Improved Debiased Machine Learning through Optimization and Calibration

Lars Wim Paul van der Laan

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington

2026

Reading Committee:
Marco Carone, Chair
Alex Luedtke, Chair
Andrea Rotnitzky

Program Authorized to Offer Degree:
Statistics

University of Washington

Abstract

Improved Debiased Machine Learning through Optimization and Calibration

Lars Wim Paul van der Laan

Chairs of the Supervisory Committee:

Marco Carone

Department of Biostatistics

Alex Luedtke

Department of Statistics

Debiased machine learning uses flexible, data-adaptive methods to estimate causal and statistical targets while correcting bias from nuisance estimation. This dissertation develops calibration- and optimization-based approaches for debiased inference and learning. The first part develops calibrated debiased machine learning for inference on linear summaries of regression functions. Standard doubly robust estimators are consistent if either the regression function or its Riesz representer is consistently estimated, but asymptotic normality typically requires both nuisances to converge sufficiently quickly. We show that calibrating nuisance estimators in AIPW procedures yields doubly robust asymptotic linearity: under partial-orthogonality conditions, these estimators are asymptotically normal when either nuisance component is estimated well, while the other may converge slowly or inconsistently. The second part introduces efficient plug-in learning for heterogeneous causal effect estimation. It constructs debiased, efficient plug-in estimators of population risks for causal contrasts, including conditional average treatment effects and relative risks, using a TMLE-style sieve correction. These estimators retain the oracle-efficiency properties of orthogonal

statistical learning while avoiding unstable pseudo-outcomes and nonconvex losses. The third part develops automatic debiased machine learning for smooth functionals of nonparametric M-estimands. This extends automatic debiasing beyond regression functions to infinite-dimensional population risk minimizers. The central object is the Hessian Riesz representer of the target derivative, which determines the leading plug-in bias and influence-function correction, reducing automatic debiasing to two risk minimization problems: one for the M-estimand and one for the representer.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am deeply grateful to my advisors, Marco Carone and Alex Luedtke, for their guidance, support, and mentorship throughout my Ph.D. Their clarity of thought, scientific standards, and generosity with their time have been instrumental in my development as a researcher. I entered graduate school with much to learn about scientific writing, and they taught me how to write with greater precision, structure, and care. I feel very fortunate to have learned from them.

I am also grateful to my father, Mark van der Laan, for introducing me to statistics, causal inference, and targeted and debiased machine learning. While I was living at home during COVID-19, before starting my Ph.D., we had countless hours-long dinner conversations about TMLE. I am grateful for his patience in answering my questions and for the wisdom he shared with me. I would also like to thank my mother and sister for their patience in letting us talk about TMLE for quite so long. His work and example have shaped the way I think about research, and his encouragement has meant a great deal to me.

I am especially grateful to my mother for her immense support throughout my Ph.D. She helped keep the rest of my life moving while I was focused on research: booking flights home and to conferences, handling reimbursements and tax returns, and taking care of countless practical details I would otherwise have neglected. Her support made the Ph.D. much easier, and I am deeply thankful for it. I would also like to thank my sister for making visits home so much fun, and especially for bringing Whitney, the newest and smallest Yorkie member of the van der Laan family, into our lives.

I am grateful to the many collaborators and mentors who shaped the work in and outside this dissertation. I thank Ahmed Alaa for invaluable mentorship on writing conference papers

and for helping me develop as an independent researcher. I thank Aurelien Bibaut for being an excellent mentor during my internship at Netflix and, later, a wonderful collaborator. I thank Nathan Kallus for his guidance and support, and for sponsoring my Ph.D. through the Netflix Graduate Research Fellowship. I thank Peter Gilbert for being a supportive mentor and for giving me the opportunity to develop causal inference methods, software, and applied biostatistical tools for biomedical research, including COVID-19 trial analyses. I am also thankful to Andrea Rotnitzky for serving on my reading committee and for her careful engagement with my work, particularly through her insightful comments during our weekly lab meetings.

I am also thankful for the friends I made during the Ph.D. program, including Nolan Cole, Charlotte Talham, Grant Hopkins, Kat Hoffman, Ellen Graham, Olivia McGough, Alex Kokot, Vydhourie Thiyageswaran, Medha Agarwal, and Ethan Ashby. Their friendship, humor, and support made graduate school much more enjoyable. I am especially grateful to Nolan for being such a supportive roommate. I will always remember café grind days with Nolan, Thursday work days at the Hutch with Charlotte, game nights with Grant, sharing Nolan's cappuccinos with Kat, reality TV nights with Ellen, and many other small moments that made graduate school fun.

Chapter 1

INTRODUCTION

1.1 Overview

This dissertation consists of three papers at the intersection of semiparametric statistics, machine learning, and causal inference. Each paper concerns inference on a target parameter that is lower-dimensional, or more structured, than the nuisance features needed to estimate it. Such targets include average and heterogeneous treatment effects, relative risks, survival summaries, and policy values. Their estimation typically depends on outcome regressions, treatment mechanisms, conditional distributions, or density ratios. The central problem is to use flexible estimates of these nuisance functions without allowing nuisance estimation error to determine the first-order behavior of the target estimator.

Flexible learning methods can reduce model bias by adapting to nonlinearities, interactions, and high-dimensional covariates. Predictive accuracy for nuisance functions, however, is not sufficient for valid inference on the target parameter. Plug-in estimators can retain first-order bias, converge at slow nonparametric rates, or yield confidence intervals with poor coverage even when the nuisance estimates perform well under predictive loss.

Debiased machine learning addresses this difficulty by combining flexible first-stage estimation with an influence-function-based correction for the leading plug-in bias. This idea appears in classical one-step estimation, estimating-equation methods, targeted minimum loss-based estimation, and double/debiased machine learning [Pfanzagl and Wefelmeyer, 1985, Bickel et al., 1993, Robins et al., 1994, 1995, van der Laan and Robins, 2003, Laan and Rubin, 2006, van der Laan and Rose, 2011, Chernozhukov et al., 2018a]. Under suitable conditions, these corrections reduce first-order sensitivity to nuisance estimation error, enabling root-n inference or oracle-type guarantees with flexible first-stage methods.

The papers in this dissertation develop calibration- and optimization-based approaches to constructing these corrections. Calibration is used to impose bias-reducing equations

on nuisance estimates before forming one-step estimators for regression functionals. Optimization is used to construct plug-in risk estimators for heterogeneous causal contrasts and to characterize debiasing directions for functionals of general nonparametric M-estimands. Together, these results extend first-order debiasing to settings in which standard DML does not fully resolve the inferential or computational problem.

1.2 Improving debiased machine learning

1.2.1 Calibration for doubly robust inference

The first paper studies doubly robust inference for regression functionals. Many causal parameters, including counterfactual means and average treatment effects, can be written as linear functionals of an outcome regression. Standard debiased machine learning estimators combine a plug-in estimate of the functional with an empirical correction involving the corresponding Riesz representer. This framework includes augmented inverse probability weighting as a special case and provides a generic construction for linear regression functionals [Bang and Robins, 2005, Newey, 1994, Chernozhukov et al., 2018b, 2022c, Rotnitzky et al., 2021, Hirshberg and Wager, 2021].

The standard DML estimator is doubly robust for consistency: it can remain consistent if either the outcome regression or the Riesz representer is consistently estimated. It also satisfies a rate double robustness property. The leading remainder is typically a product of the two nuisance errors, so asymptotic linearity follows when this product is $o_p(n^{-1/2})$ [Bradic et al., 2019a, Smucler et al., 2019, Rotnitzky et al., 2021]. This requirement is much weaker than requiring both nuisance estimators to converge at parametric rates. It is not, however, doubly robust inference. If one nuisance estimator is inconsistent, or converges too slowly, the product remainder may be first order and standard Wald confidence intervals can be invalid, even when the point estimator remains consistent.

Several approaches address this limitation. Sparsity-based methods obtain asymptotic normality under one-sided correctness in high-dimensional models by using regularization-induced moment conditions [Athey et al., 2018, Bradic et al., 2019b,a, Smucler et al., 2019, Tan, 2020, Liu et al., 2023]. These guarantees are powerful but depend on sparsity and

on particular nuisance estimators. A separate line of work develops targeted minimum loss-based estimators for doubly robust inference, beginning with the average treatment effect and extending to missing outcomes, survival curves, and partially linear models [van der Laan, 2014, Benkeser et al., 2017, Díaz and van der Laan, 2017, Díaz, 2019, Dukes et al., 2024]. These methods are learner-agnostic and can deliver doubly robust asymptotic linearity, but they rely on iterative targeting steps and additional nuisance regressions. New parameters also tend to require new targeting constructions and new remainder analyses.

Calibration provides another route to doubly robust inference. In predictive modeling, calibration adjusts fitted values so that they are self-consistent with the targets they are meant to predict, often through one-dimensional methods such as isotonic regression [Zadrozny and Elkan, 2001, 2002, Niculescu-Mizil and Caruana, 2005b]. For squared-error regression, this self-consistency can be expressed as empirical orthogonality of residuals against functions of the fitted values. More generally, calibration with respect to the nuisance loss enforces analogous empirical score equations. The key observation is that these equations can be chosen to impose the low-dimensional orthogonality conditions needed to remove the leading cross-product remainder of a DML estimator.

Chapter 2 develops this calibration-based approach. The chapter shows that calibrating the outcome-regression and Riesz-representer estimates before forming the one-step estimator enforces the empirical orthogonality equations needed to debias the cross-product remainder. The proposed implementation uses a post-hoc, one-dimensional isotonic regression adjustment applied to cross-fitted nuisance estimates. Under a partial orthogonality condition and regularity conditions, the calibrated DML estimator is asymptotically normal when either nuisance component is estimated sufficiently well, while the other may converge slowly or even inconsistently. Compared with TMLE-style approaches, the calibration step is non-iterative, applies to a class of linear regression functionals, and supports bootstrap-assisted confidence intervals without estimating the additional nuisance functions that appear in the misspecified limiting distribution.

1.2.2 Efficient plug-in risk minimization for heterogeneous causal learning

The second paper studies learning causal contrasts that are themselves functions, such as the conditional average treatment effect and the conditional relative risk. A simple approach is plug-in learning, or T-learning: estimate the relevant outcome regressions and substitute them into the contrast [Künzel et al., 2019]. Plug-in learners are easy to combine with standard supervised learning methods and often preserve useful structure, such as boundedness of the contrast or convexity of the risk. Their limitation is statistical efficiency. The outcome regressions may be more complex than the contrast, so estimating the full regression surface can be unnecessarily difficult when the scientific target is a lower-complexity function [Kennedy, 2023a, Nie and Wager, 2021].

Orthogonal statistical learning addresses this limitation by constructing risks or losses whose first-order sensitivity to nuisance estimation error is removed by Neyman orthogonality [Rubin and van der Laan, 2006, Kennedy, 2023a, Nie and Wager, 2021, Foster and Syrgkanis, 2023, Morzywolek et al., 2023]. The DR-learner and R-learner are prominent examples for the conditional average treatment effect. These methods often use AIPW-style pseudo-outcomes or residualized losses, and can achieve oracle-type guarantees under suitable conditions. The cost is that the empirical criterion may be less stable than the original plug-in risk. It can involve inverse propensity weights, extreme pseudo-outcomes, or nonconvex losses even when the target risk is convex. For conditional relative-risk learning, the orthogonalized logistic-type risk can involve negative weights and pseudo-outcomes outside the unit interval, making standard optimization tools difficult to use [Jiang et al., 2019, Qiu et al., 2019].

Chapter 3 develops a plug-in alternative to explicit orthogonalization. The starting point is that many orthogonalized risk estimators can be written as a plug-in risk estimator plus an augmentation term. Rather than adding this term to the empirical criterion, the chapter modifies the nuisance estimate so that the augmentation term is asymptotically negligible. This is the risk-minimization analogue of targeted minimum loss-based estimation: the final estimator remains a plug-in estimator, but the nuisance estimate is corrected to solve the relevant score equations [Laan and Rubin, 2006, van der Laan and Rose, 2011]. Because

the population risk is indexed by the candidate contrast function, this correction must hold uniformly over the function class. The chapter achieves this with a sieve-based targeting step that makes the plug-in risk asymptotically equivalent to an efficient one-step risk estimator.

The resulting learner retains the stability of plug-in optimization while recovering the large-sample behavior of orthogonal learning. For the conditional average treatment effect, it replaces AIPW-style pseudo-outcomes with a plug-in contrast formed from a debiased outcome regression. For the conditional relative risk, it yields a convex weighted logistic-regression problem with nonnegative weights and valid pseudo-outcomes, avoiding the nonconvexity of the explicit one-step risk. Under regularity conditions, the efficient plug-in learner is asymptotically equivalent to the minimizer of an oracle-efficient one-step risk estimator.

1.2.3 Loss-based automatic debiasing for M-estimands

The third theme is automatic debiasing. In many semiparametric problems, the difficult step is not only estimating nuisance functions, but deriving the correction itself. Classical one-step, estimating-equation, and targeted learning estimators require an influence-function representation of the target parameter [Bickel et al., 1993, van der Laan and Robins, 2003, van der Laan and Rose, 2011, Chernozhukov et al., 2018a]. Deriving this representation can be technically demanding and parameter-specific, especially when the target depends on nuisance functions defined implicitly through optimization problems.

Automatic DML simplifies this task for an important class of problems. For linear functionals of regression functions, the influence-function correction is determined by the Riesz representer of the target functional, and this representer can itself be estimated by solving an auxiliary risk minimization problem [Newey, 1994, Chernozhukov et al., 2021b, 2022c, Williams et al., 2025]. This makes debiasing modular: after specifying the regression loss and the target functional, the analyst estimates both the regression function and its Riesz representer using supervised learning tools. Direct representer estimation can also be preferable to estimating a full collection of primitive nuisance functions, such as propensity scores or density ratios, and then algebraically combining them.

Existing automatic debiasing methods remain largely organized around regression functionals and related extensions, including covariate shift, sequential regressions, and certain inverse problems [Chernozhukov et al., 2022c, 2023, Bennett et al., 2023b, 2025, van der Laan et al., 2025c]. These extensions share a common structure, but the representers, nuisance components, and debiasing arguments are often developed separately. At the same time, many modern targets are functionals of objects that are not ordinary regression functions. They may depend on quantile or survival functions, density ratios, pseudo-outcome regressions, causal contrast functions, or solutions to ill-posed inverse problems [Hines and Miles, 2025, Wang et al., 2018, Westling et al., 2024, Yang et al., 2023, Foster and Syrgkanis, 2023, Carrasco et al., 2007]. These objects are naturally defined as nonparametric M-estimands: minimizers of population risks over infinite-dimensional spaces.

Chapter 4 extends automatic DML to smooth functionals of such nonparametric M-estimands. The key object is the Riesz representer of the target derivative under the Hessian inner product induced by the population risk. This Hessian Riesz representer characterizes the leading plug-in bias and determines the loss-gradient correction in the efficient influence function. As a result, debiasing reduces to two optimization problems: one for the M-estimand and one for its representer. The chapter develops the corresponding von Mises expansion, efficient influence-function representation, one-step estimator, targeted minimum loss-based estimator, and sieve plug-in estimator. For quadratic risks and linear targets, the resulting estimators satisfy doubly robust product-rate conditions in the M-estimand and Hessian Riesz representer.

1.3 Organization

The remainder of the dissertation follows the three-paper structure. Chapter 2 develops calibrated debiased machine learning for doubly robust inference on linear functionals of regression functions. Chapter 3 develops efficient plug-in learning for heterogeneous causal contrasts, including the conditional average treatment effect and conditional relative risk. Chapter 4 develops automatic debiased machine learning for smooth functionals of nonparametric M-estimands.

Chapter 2

DOUBLY ROBUST INFERENCE VIA CALIBRATION

Lars van der Laan, Alex Luedtke, and Marco Carone

Abstract

Doubly robust estimators are widely used for estimating average treatment effects and other linear summaries of regression functions. While consistency requires only one of two nuisance functions to be estimated consistently, asymptotic normality for linear functionals typically requires sufficiently fast convergence of both. We address this mismatch by showing that calibrating the nuisance estimators within a doubly robust procedure yields doubly robust asymptotic normality, provided that one nuisance estimator suitably captures the estimation error in the other. We introduce a general framework, *calibrated debiased machine learning* (calibrated DML), and propose a specific estimator that augments standard DML with a simple isotonic regression adjustment. Under a partial orthogonality condition, our theoretical analysis shows that the calibrated DML estimator remains asymptotically normal if either the regression or the Riesz representer of the functional is estimated sufficiently well, allowing the other to converge arbitrarily slowly or even inconsistently. We further propose a simple bootstrap method for constructing confidence intervals, enabling doubly robust inference without additional nuisance estimation. In a range of semi-synthetic benchmark datasets, calibrated DML reduces bias and improves coverage relative to standard DML. Our method can be integrated into existing DML pipelines by adding just a few lines of code to calibrate cross-fitted estimates via isotonic regression.

2.1 Introduction

2.1.1 Motivation

Nonparametric inference on the causal effects of dynamic and stochastic treatment interventions using flexible statistical learning techniques is a fundamental problem in causal inference. These inferential targets fall within a broader class of target parameters that can be expressed as linear functionals of the outcome regression, that is, the conditional mean of the outcome given treatment assignment and baseline covariates [Chernozhukov et al., 2018b, Rotnitzky et al., 2021, Hirshberg and Wager, 2021]. To achieve inference for such parameters, various debiased machine learning (DML) frameworks are available, including one-step estimation [Pfanzagl and Wefelmeyer, 1985, Bickel et al., 1993], estimating equations and double machine learning [Robins et al., 1994, 1995, 2000, van der Laan and Robins, 2003, Chernozhukov et al., 2018a], and targeted minimum loss estimation (TMLE) [Laan and Rubin, 2006, van der Laan and Rose, 2011], and sieve-based estimators [Qiu et al., 2021, Cho et al., 2024]. When performing inference on linear functionals, debiased machine learning methods frequently produce estimators that demonstrate robustness against inconsistent or slow estimation of certain nuisance functions, a property known as double robustness [Bang and Robins, 2005, Smucler et al., 2019]. For example, when estimating the average treatment effect (ATE), doubly robust estimators remain consistent as long as either the outcome regression or the propensity score is consistently estimated, even if the other is inconsistently estimated. This robustness property extends more generally to linear functionals of the outcome regression, necessitating consistent estimation of only the outcome regression itself or the Riesz representer of the linear functional [Chernozhukov et al., 2018b, Rotnitzky et al., 2021].

Extending the double robustness property of debiased machine learning estimators to statistical inference is challenging. By *double robustness for statistical inference*, we mean that valid root-n inference remains possible when one nuisance function is estimated consistently at a sufficiently fast rate, even if the other is not. In standard debiased machine learning procedures, however, asymptotic normality typically requires both the outcome regression and the Riesz representer to be estimated consistently and sufficiently quickly, so that the

product of their estimation errors is $o_p(n^{-1/2})$. If this condition fails, desirable properties such as asymptotic linearity, asymptotic normality, and $n^{1/2}$-consistency can be lost, compromising the validity of inference [Benkeser et al., 2017]. In such cases, standard confidence intervals and p-values can be invalid. In practice, inconsistency may arise from model misspecification, such as omitted interactions or nonlinearities, or from over-regularization of flexible learners. Slow convergence may arise even under consistency when the nuisance functions are intrinsically complex, for example because they involve nonsparse high-dimensional interactions or have limited smoothness. One approach to achieving doubly robust inference is to use parametric methods, such as maximum likelihood estimation, to estimate the nuisance parameters. These methods exhibit model double robustness, maintaining asymptotic linearity when at least one of the nuisance parametric models is correctly specified [Vermeulen and Vansteelandt, 2015, 2016, Shook-Sa et al., 2025]. However, their usefulness for nonparametric inference is limited by their reliance on strong modeling assumptions, since in practice both models are often misspecified [Kang and Schafer, 2007].

To overcome the limitations of parametric methods, several estimators exhibiting *sparsity double robustness*—a weaker form of model double robustness—have been proposed [Avagyan and Vansteelandt, 2017, Athey et al., 2018, Bradic et al., 2019b,a, Smucler et al., 2019, Tan, 2020, Zhang et al., 2021, Ghosh and Tan, 2022, Bradic et al., 2024, Baybutt and Navjeevan, 2026]. In high-dimensional generalized linear models, Avagyan and Vansteelandt [2017] and Bradic et al. [2019b,a] introduced doubly robust ℓ_1–regularized estimators of the ATE that are asymptotically normal when either the outcome regression or propensity score is sufficiently sparse, allowing misspecification of one nuisance model. Relatedly, Athey et al. [2018] and Tan [2020] proposed singly robust estimators that are robust to misspecification of the outcome regression, assuming sufficient sparsity of the propensity score. A key limitation of these approaches is their reliance on strong sparsity assumptions, which may fail in practice. Bradic et al. [2019a] study doubly robust inference under approximate sparsity, but still impose a product-rate condition on the approximation errors of both nuisance functions with respect to the chosen basis dictionary (Assumption 2). Smucler et al. [2019] develop a unifying ℓ_1-regularization framework for sparse doubly robust estimation, accommodating model misspecification by allowing the nuisance estimators to converge to

approximately sparse, potentially misspecified limits at a sufficiently fast rate. For the ATE, Liu et al. [2023] relax sparsity requirements on misspecified limits by assuming sparsity only for the consistently estimated nuisance, albeit at the cost of irregularity and loss of efficiency. Finally, these methods are not learner-agnostic: they are largely tied to ℓ_1-regularized generalized linear models for the nuisance functions, which limits their ability to leverage modern machine-learning methods.

By contrast, the line of work initiated by van der Laan [2014] and further developed by Benkeser et al. [2017] provides a learner-agnostic debiasing framework—based on targeted maximum likelihood estimation (TMLE) [van der Laan and Rose, 2011]—for doubly robust inference with nuisance functions estimated using generic machine learning methods. Under suitable conditions, the resulting estimators are *doubly robust asymptotically linear* (DRAL): they remain asymptotically linear, and hence asymptotically normal under standard conditions, even if one nuisance estimator is inconsistent, provided the other is estimated sufficiently well (i.e., at rate $o_p(n^{-1/4})$). Concretely, for the ATE, if π_n and μ_n estimate the propensity score π_0 and outcome regression μ_0, then a DRAL estimator can remain asymptotically linear under the *best-rate* condition that $\|\pi_n - \pi_0\| = o_p(n^{-1/4})$ or $\|\mu_n - \mu_0\| = o_p(n^{-1/4})$, whereas rate double robustness imposes the more stringent product-rate condition $\|\mu_n - \mu_0\| \, \|\pi_n - \pi_0\| = o_p(n^{-1/2})$. For the ATE, these methods iteratively update initial propensity score and outcome regression estimates using low-dimensional (often one-dimensional) regressions that enforce moment conditions and debias the cross-product remainder in the asymptotic expansion. This approach has since been extended to nonparametric inference for mean outcomes under informative missingness [Díaz and van der Laan, 2017], counterfactual survival curves [Díaz, 2019], and semiparametric inference for the ATE in partially linear regression models [Dukes et al., 2024].

The TMLE-based framework for doubly robust inference has some limitations. First, the iterative debiasing procedure is difficult to study theoretically and can be computationally demanding to implement, posing challenges for its application to large-scale data or complex data structures. While Bonvini et al. [2024] proposed a non-iterative, nonparametric DRAL estimator of the ATE based on higher-order influence functions, their procedure requires careful tuning of bivariate kernel-smoothing parameters, for which practical guidance

Figure 2.1: We show **calibrating the nuisances** before debiasing ensures doubly robust asymptotic normality.

remains limited. Second, when considering new parameters of interest, a new debiasing algorithm must be derived on a case-by-case basis, requiring careful analysis of the relevant remainder terms and novel targeting constructions. To date, there is no unified, learner-agnostic approach to doubly robust inference for a generic linear functional of the outcome regression. Finally, the available guarantees are mainly designed for one-sided misspecification and do not strengthen inference when both nuisance estimators are consistent but one converges too slowly for the product-rate condition to hold. Indeed, they typically assume that both nuisance estimators converge to their (possibly misspecified) limits at rate $o_p(n^{-1/4})$.

2.1.2 Contributions of this work

We develop new methods for doubly robust inference on linear functionals of the outcome regression. Our goal is to construct DRAL estimators that permit valid inference when one of the two nuisance functions is estimated arbitrarily slowly or even inconsistently, provided that at least one nuisance converges at an $o_p(n^{-1/4})$ rate. We first establish a link between calibration, a technique typically used in prediction and classification tasks [Zadrozny and Elkan, 2001, Niculescu-Mizil and Caruana, 2005b], and doubly robust asymptotic linearity. We then introduce a general framework, *calibrated debiased machine learning* (calibrated DML). As outlined in Figure 2.1, calibrated DML begins with standard cross-fitted estimates of the outcome regression and Riesz representer, obtained using any machine learning algorithm. It then adds a single post-hoc step: each nuisance estimator is *calibrated* by fitting a one-dimensional map from its fitted values to its target using, e.g., isotonic regression [Barlow and Brunk, 1972]. We show that this calibration step enforces an appropriate empirical orthogonality between the nuisance errors, which enables a one-step debiased estimator

Table 2.1: Properties of G-computation plug-in, IPW, AIPW and calibrated DML (proposed approach) estimators of the ATE when nuisances are estimated using machine learning, in settings where only the propensity score vs. only the outcome regression vs. both nuisances are estimated well (i.e., consistently and at a rate faster than $n^{-1/4}$). Report whether each estimator is **consistent**, asymptotically **normal** at root-sample size rate, and nonparametric **efficient**.
(∗) Our estimator is efficient if one nuisance is estimated well), provided the other is estimated consistently (at any rate).

	only propensity score estimated well			only outcome regression estimated well			both nuisance functions estimated well		
estimator	**consistent**	**normal**	**efficient**	**consistent**	**normal**	**efficient**	**consistent**	**normal**	**efficient**
G-computation	–	–	–	✓	×	×	–	–	–
IPW	✓	×	×	–	–	–	–	–	–
AIPW	✓	×	×	✓	×	×	✓	✓	✓
Calibrated DML (ours)	✓	✓	✓∗	✓	✓	✓∗	✓	✓	✓

to achieve doubly robust asymptotic linearity. This is achieved under a partial orthogonality condition comparable to those assumed in prior work on learner-agnostic doubly robust inference, including van der Laan [2014], Benkeser et al. [2017], Díaz and van der Laan [2017], Dukes et al. [2024], Bonvini et al. [2024]. Confidence intervals can be constructed using the closed-form influence function of the calibrated DML estimator, without requiring knowledge of which nuisance estimators converge sufficiently quickly. We show that no correction is needed to the standard influence-function-based variance estimator when both nuisance estimators are consistent for the true nuisances, even if one converges arbitrarily slowly. When one nuisance may be inconsistently estimated, we propose a computationally efficient bootstrap-assisted procedure that avoids the additional nuisance estimation required in van der Laan [2014] and related approaches, and remains valid under the same nuisance-rate conditions.

Table 2.1 summarizes the properties of a specific implementation of calibrated DML in the well-studied setting of ATE estimation. This implementation employs isotonic regression for calibration and a cross-fitted augmented inverse probability weighted (AIPW) estimator for debiasing. Compared to the standard AIPW estimator, calibrated DML remains asymptotically normal under weaker convergence requirements: it suffices that one nuisance be consistently estimated at an $o_p(n^{-1/4})$ rate (the other may be inconsistent). It is efficient when *both* nuisances are consistently estimated (with at least one attaining an $o_p(n^{-1/4})$ rate). Calibrated DML can yield even larger improvements over non-doubly robust esti-

mators based on the G-computation formula or inverse probability weighting (IPW), since these estimators are neither doubly robust nor generally asymptotically normal under flexible nuisance estimation.

Our approach is motivated by the DRAL estimators of the ATE proposed by van der Laan [2014] and developed further in subsequent work [Benkeser et al., 2017, Díaz and van der Laan, 2017, Díaz, 2019, Dukes et al., 2024]. We share the objective of constructing nuisance estimators that achieve doubly robust asymptotic linearity by solving certain score equations, but our strategy and implementation differ substantially, even for the ATE. In particular, our approach is a non-iterative and computationally efficient two-step procedure that can be implemented using commonly available isotonic regression software. Moreover, our method does not depend on the specific form of the linear functional; rather, it requires initial nuisance estimators and black-box access to an empirical estimator of the functional. We share with Benkeser et al. [2017] the property that our estimator admits Wald-type inference based on a closed-form influence function, without requiring the analyst to know which nuisance is correctly specified. Unlike Benkeser et al. [2017], however, our proposed bootstrap-assisted inference procedure avoids estimating the additional nuisance components that appear in the closed-form influence function, which can simplify implementation and potentially improve finite-sample performance. Finally, our theoretical results refine those of prior work by showing that our estimator achieves asymptotic linearity not only when one nuisance function is consistently estimated, but also when both nuisance functions are consistently estimated and one converges too slowly. Another distinction is that we consider doubly robust inference for generic continuous linear functionals, which requires a new analysis of the doubly robust remainder that leverages the Riesz representation property.

2.1.3 Related work

Calibration in causal inference. Prior work has studied the finite-sample benefits of calibrating propensity score estimates in IPW and AIPW estimators of the ATE [Gutman et al., 2022, Deshpande and Kuleshov, 2023, Ballinari and Bearth, 2024, van der Laan et al., 2024b, Rabenseifner et al., 2025]. These studies focus primarily on calibrating a single nui-

sance function—the propensity score—for a single parameter (the ATE), and they do not provide theoretical justification beyond the observation that calibration can reduce nuisance estimation error. We extend this line of work by showing that, for general linear functionals, calibrating *both* the outcome regression and the Riesz representer is key for further debiasing (in finite samples and asymptotically) and for constructing DRAL estimators. More broadly, our work shows that calibration can yield asymptotic guarantees and relax nuisance-rate requirements, rather than serving only to improve the finite-sample quality of nuisance estimators. In the special case of the ATE, we find that calibrating the outcome regression can reduce finite-sample bias and remove certain asymptotic bias terms.

Sparsity double robustness. Our calibration-based approach to doubly robust inference is closely related to the literature on (approximate) sparsity-based doubly robust inference discussed in the Introduction [Athey et al., 2018, Bradic et al., 2019a, Smucler et al., 2019, Tan, 2020, Liu et al., 2023]. Methodologically, our approach enforces one-dimensional orthogonality conditions on nuisance errors via calibration (see (2.4)–(2.5)), whereas sparsity-based methods typically enforce high-dimensional moment conditions. In particular, when the nuisances are estimated by ℓ_1-regularized empirical risk minimization over linear classes, the Karush–Kuhn–Tucker (KKT) conditions yield approximate empirical orthogonality over suitable linear subspaces (see, e.g., Tan, 2020, Eq. 15; Bradic et al., 2019b, Sec. 2.1; Smucler et al., 2019, Sec. 4.1). We also note a terminology difference: Tan [2020], Liu et al. [2023], and related work use "calibration" in the survey-sampling sense, referring to adjustments that enforce moment conditions [Deville and Särndal, 1992], whereas we use "calibration" in the machine-learning sense to mean mean-unbiasedness conditional on the model's own prediction, a one-dimensional objective. The analogous high-dimensional goal is closely related to *multi-accuracy*; see Roth, 2022. We discuss these differences in more detail in Appendix 2.A.

Our closest comparison is Liu et al. [2023]: both methods apply a post-hoc correction, but Liu et al. [2023] solve a high-dimensional moment-matching problem built on preliminary ℓ_1-regularized generalized linear models, whereas we use a one-dimensional calibration step and allow arbitrary user-supplied first-stage learners. The guarantees differ accordingly:

we show that one-dimensional calibration suffices for DRAL under a partial-orthogonality condition (the upcoming B1) on the nuisances, whereas Liu et al. [2023] rely on sparsity conditions tailored to their high-dimensional setting and require an auxiliary sample (e.g., via sample splitting) and careful tuning. Efficiency guarantees also differ: our calibrated DML estimators are semiparametrically efficient when both nuisances are consistent and at least one converges faster than $n^{-1/4}$ (Table 2.1), whereas Liu et al. [2023] do not generally establish semiparametric efficiency. In contrast, related works, such as Bradic et al. [2019a] and Smucler et al. [2019], establish efficiency under additional approximate sparsity assumptions on the (potentially misspecified) limit of the slower nuisance estimator.

Higher-order influence functions and undersmoothing. Higher-order influence function (HOIF) estimators [Robins et al., 2008, Liu et al., 2017] use higher-order bias corrections to relax the nuisance-rate requirements of first-order doubly robust methods, and can yield asymptotically valid and efficient inference even when nuisance estimators fail to satisfy the usual $o_p(n^{-1/2})$ product-rate condition. Related robustness to slow nuisance rates can also be obtained without explicit HOIF-based corrections by combining sample splitting with undersmoothing [Newey and Robins, 2018, McGrath and Mukherjee, 2022, McClean et al., 2024]. While HOIF methods can accommodate flexible first-stage machine-learning estimators, their higher-order bias corrections typically rely on smoothness and approximability assumptions, often of Hölder type, to control higher-order remainders. They also introduce additional tuning parameters—such as the choice of basis functions, truncation order, sieve dimension, or bandwidth—for which practical guidance can be limited. Consequently, HOIF estimators can be more complex to implement and tune, and their performance may be sensitive to these assumptions and choices, especially in high-dimensional settings.

For the expected conditional covariance parameter, Bonvini et al. [2024] show that modified HOIF-style estimators can attain guarantees similar to those of the TMLE-style doubly robust inference approach of Benkeser et al. [2017]. Their HOIF-based DRAL estimators use the outcome regression and propensity score as a bivariate dimension reduction and achieve validity under conditions comparable to those imposed in our work and related TMLE-based approaches [Benkeser et al., 2017]. In contrast, calibrated DML uses a simple

post-hoc calibration step to control the relevant higher-order bias, requiring no additional nuisance estimation beyond standard first-order methods. When implemented with isotonic calibration, this post-hoc step is also tuning-free.

Organization. This paper is organized as follows. In Section 2.2, we introduce the class of regression functionals we consider and provide background on a standard rate-doubly-robust debiased estimator. In Section 2.3, we formulate the problem of doubly robust inference, introduce a model-agnostic linear expansion for the bias of one-step debiased estimators, and demonstrate how further debiasing of nuisances can be employed to construct DRAL estimators. In Section 2.4, we present a novel link between nuisance estimator calibration and the doubly robust inference properties of debiased machine learning estimators, use this link to build our general calibrated DML framework, and propose a calibrated DML estimator based on isotonic regression with accompanying bootstrap-based confidence intervals. In Section 2.5, we establish the theoretical properties of our proposed calibrated DML estimators and confidence intervals. Finally, we empirically investigate the performance of calibrated DML in Section 2.6 and provide concluding remarks in Section 2.7.

2.2 Problem setup

2.2.1 Data structure and target parameter

In this section, we introduce the data structure and the class of regression functionals under consideration, along with their semiparametric efficiency bound. In the next subsection, we review the standard one-step debiased estimator and its rate double robustness property. We then explain why rate double robustness alone is insufficient for doubly robust inference, thereby motivating our objective of doubly robust asymptotic linearity in Section 2.3.

Suppose we observe an i.i.d. sample $Z_1, \ldots, Z_n$ of $Z = (W, A, Y) \sim P_0$, where $W \in \mathcal{W} \subset \mathbb{R}^d$ are baseline covariates, $A \in \mathcal{A} \subset \mathbb{R}^p$ is a treatment, and $Y \in \mathbb{R}$ is a bounded outcome. For any distribution P, let $\mu_P(a, w) := E_P\{Y \mid A = a, W = w\}$ denote the outcome regression and $Pf := \int f(z)\, dP(z)$ for any integrable function f. We denote by $P_{0,A,W}$ the marginal distribution of (A, W) under P_0, by $\mathcal{H} := L^2(P_{0,A,W})$ the corresponding Hilbert space of square-integrable functions on $\mathcal{A} \times \mathcal{W}$ with inner product

$\langle \mu_1, \mu_2 \rangle := \int \mu_1(a,w)\mu_2(a,w)\, P_{0,A,W}(da,dw)$, and by $\|\cdot\|$ the induced norm. For general P, we write $\|\mu\|_P := \{\int \mu(a,w)^2\, P_{A,W}(da,dw)\}^{1/2}$, where $P_{A,W}$ is the marginal law of (A,W) under P. Let $\mathcal{M}$ be a nonparametric model containing P_0. To simplify notation, we write S_0 for any summary S_{P_0} of the true distribution P_0.

Our objective is to develop doubly robust nonparametric inference for a linear summary $\tau_0 := \psi_0(\mu_0)$ of the outcome regression μ_0, where $\psi_0 : \mathcal{H} \to \mathbb{R}$ is an L^2-continuous linear functional. We allow the functional of interest to depend on the unknown data-generating distribution P_0, viewing ψ_0 as the evaluation at P_0 of a smooth mapping $P \mapsto \psi_P$ defined on the statistical model $\mathcal{M}$. As illustrated at the end of this subsection, many parameters take this form, including average causal effects; this framework also subsumes the class of continuous linear functionals studied in Chernozhukov et al. [2022c], Bradic et al. [2019a].

We impose the following smoothness conditions on the mapping $(P,\mu) \mapsto \psi_P(\mu)$, which are sufficient for pathwise differentiability of the composite parameter $\Psi : P \mapsto \psi_P(\mu_P)$ at P_0. For each $P \in \mathcal{M}$, we view ψ_P as a bounded linear functional on the normed space $(\mathcal{H}, \|\cdot\|_{L^2(P_{0,A,W})})$, equipped with the corresponding operator norm.

A1) *(linearity and boundedness)* it holds that $\psi_0(c\mu_1 + \mu_2) = c\psi_0(\mu_1) + \psi_0(\mu_2)$ for each $\mu_1, \mu_2 \in \mathcal{H}$ and $c \in \mathbb{R}$ and that $\sup_{\mu \in \mathcal{H}\setminus\{0\}} |\psi_0(\mu)|/\|\mu\| < \infty$;

A2) *(Hellinger continuity)* $P \mapsto \psi_P$ is a continuous map at P_0 with domain $\mathcal{M}$ equipped with the Hellinger distance and codomain $\{\psi_P : P \in \mathcal{M}\}$ equipped with the operator norm;

A3) *(pathwise differentiability)* for each $\mu \in \mathcal{H}$, $P \mapsto \psi_P(\mu)$ is pathwise differentiable at P_0 relative to $\mathcal{M}$ with efficient influence function $\phi_{0,\mu} \in L_0^2(P_0)$.

A key object associated with a continuous linear functional is its Riesz representer, which plays a central role in debiasing and in the construction of asymptotically linear estimators of τ_0 (see Williams et al. [2025] for an accessible introduction). Since ψ_0 is a continuous linear functional on $\mathcal{H}$, the Riesz representation theorem guarantees the existence of a unique $\alpha_0 \in \mathcal{H}$ such that

$$\psi_0(\mu) = \langle \alpha_0, \mu \rangle \quad \text{for all } \mu \in \mathcal{H}. \tag{2.1}$$

Section 2.2.2 shows how (2.1) motivates automatic debiased machine learning estimators of τ_0. Under our conditions, the same α_0 also appears in the nonparametric efficient influence function (EIF) for the parameter mapping Ψ, and hence in the associated semiparametric efficiency bound [Bickel et al., 1993]. Letting $D_{\mu,\alpha}$ denote the map $z \mapsto \alpha(a,w)\{y-\mu(a,w)\}$, the following theorem identifies the EIF.

Theorem 2.1 (Pathwise differentiability)**.** *Under A1–A3, the parameter $\Psi : \mathcal{M} \to \mathbb{R}$ is pathwise differentiable at P_0 and has nonparametric efficient influence function $D_0 := \phi_{0,\mu_0} + D_{\mu_0,\alpha_0}$, where ϕ_{0,μ_0} (as defined in A3) is the component arising from the dependence of ψ_0 on P_0.*

Following Van der Vaart [2000], we call an estimator (asymptotically) efficient for τ_0 (with respect to the nonparametric model $\mathcal{M}$) if it is asymptotically linear with influence function equal to the efficient influence function D_0. By the convolution theorem (e.g., Van der Vaart, 2000, Ch. 25; see also Van Der Vaart and Wellner, 1996), any regular estimator of τ_0 has asymptotic variance at least $P_0 D_0^2$, and this lower bound is attained by efficient estimators. Moreover, efficient estimators are locally asymptotically minimax among all estimators [Hájek, 1972, Le Cam, 2012].

A key class of parameters covered by our framework consists of linear functionals that can be represented as expectations of a possibly unknown linear map.

Example 2.2 (Expectation-representable linear functionals)**.** Many parameters of interest in causal inference, such as counterfactual means and the ATE, can be written as functionals of the form [Chernozhukov et al., 2022c, Bradic et al., 2019a, Rotnitzky et al., 2021]

$$\mu \mapsto \psi_0(\mu) = E_0\{m(Z,\mu)\}, \tag{2.2}$$

for a known map $m : \mathcal{Z} \times \mathcal{H} \to \mathbb{R}$. For example, the counterfactual mean outcome under a dynamic intervention that sets treatment to $d(w) \in \mathcal{A}$ for individuals with covariate value $w \in \mathcal{W}$ is obtained by taking $m : (z,\mu) \mapsto \mu(d(w),w)$, whereas the ATE corresponds to $m : (z,\mu) \mapsto \mu(1,w) - \mu(0,w)$. Doubly robust inference for the ATE parameter was considered in van der Laan [2014] and Benkeser et al. [2017]

For such functionals, ψ_0 satisfies Condition A1 if m is linear in its last argument and there exists $c < \infty$ such that, for each $\mu \in \mathcal{H}$, $\|m(\cdot,\mu)\| \leq c\|\mu\|$. Condition A2 holds by

uniform Hellinger continuity of the expectation map $P \mapsto E_P[m(Z,\mu)]$ (see Lemma 2.6 in Appendix 2.B). Finally, Condition A3 holds with $\phi_{0,\mu} := m(\cdot,\mu) - \psi_0(\mu)$. More generally, our framework also allows for nuisance-dependent functionals of the form $\psi_0(\mu) = E_0\{m_0(Z,\mu)\}$, where m_0 is an unknown map that may depend on P_0. One such case is the causal effect under a stochastic intervention, for which $m_0(z,\mu) := \int \mu(a,w)\,\omega_0(a \mid w)\,da$ with an unknown conditional density ω_0. □

In Appendix 2.F.1, we show that the expected conditional covariance parameter arising in semiparametric ATE estimation under partially linear models [Robinson, 1988, Li et al., 2011], as a measure of "causal influence" [Díaz, 2024], and in conditional independence testing [Shah and Peters, 2020] also falls within our framework; doubly robust inference for this parameter was previously studied by Dukes et al. [2024] and Bonvini et al. [2024]. We further show that the average treatment effect under covariate-dependent outcome missingness likewise fits within our framework; doubly robust inference for this estimand was previously considered by Díaz and van der Laan [2017].

2.2.2 One-step debiased estimation and rate double robustness

Suppose we have an estimator μ_n of the outcome regression μ_0 and an estimator ψ_n of the target functional ψ_0. For the expectation-representable functionals described in Example 2.2, a natural choice is the empirical plug-in map

$$\psi_n(\mu) := \frac{1}{n}\sum_{i=1}^{n} m(Z_i,\mu).$$

More generally, our theory allows nuisance-dependent functionals, provided ψ_n satisfies Condition C4 in Section 2.5. A natural estimator of the target parameter τ_0 is then the plug-in estimator $\psi_n(\mu_n)$. However, when μ_n is obtained using flexible statistical learning methods, $\psi_n(\mu_n)$ is typically not $n^{1/2}$-consistent and may fail to be asymptotically normal, precluding valid inference. The issue is excess bias: the error $\psi_0(\mu_n) - \psi_0(\mu_0)$ typically depends linearly on $\mu_n - \mu_0$ [van der Laan and Rose, 2011]. Debiasing methods aim to remove this leading bias and thereby restore root-n inference.

A widely used approach is *automatic debiasing* [Chernozhukov et al., 2018b, 2022c, van der Laan et al., 2025b], which augments the plug-in estimator with a bias correction

based on an estimate of the Riesz representer α_0 of ψ_0. Indeed, by the Riesz representation property (2.1),

$$\psi_0(\mu_n) - \psi_0(\mu_0) = \langle \alpha_0, \mu_n - \mu_0 \rangle,$$

so the leading plug-in error is linear in $\mu_n - \mu_0$. Given an estimator α_n of α_0, this expression motivates the *one-step debiased estimator*

$$\tau_n := \psi_n(\mu_n) + \frac{1}{n}\sum_{i=1}^{n} \alpha_n(A_i, W_i)\{Y_i - \mu_n(A_i, W_i)\},$$

where the second term provides an empirical bias correction. Equivalently, τ_n can be derived from an efficient influence-function adjustment (Theorem 2.1) [Bickel et al., 1993]. The nuisance functions μ_n and α_n may be learned using flexible supervised learning methods. For example, μ_n can be fit by least squares, $(z, \mu) \mapsto \{y - \mu(a, w)\}^2$, while α_n can be obtained by minimizing the empirical Riesz loss, $(z, \alpha) \mapsto \alpha^2(a, w) - 2\psi_n(\alpha)$ [Chernozhukov et al., 2018b, 2022a]. The estimator τ_n generalizes the augmented inverse probability weighted estimator [Robins et al., 2000, 1995, Bang and Robins, 2005, Bruns-Smith et al., 2025] and is also known as the automatic DML (autoDML) estimator [Chernozhukov et al., 2022c,a, van der Laan et al., 2025b]. Here, "automatic" refers to the fact that both the form of τ_n and the losses used to learn μ_n and α_n are generic, and do not require tailoring to the specific functional ψ_n.

A key property of the one-step debiased estimator τ_n is *rate double robustness* [Bradic et al., 2019a, Rotnitzky et al., 2021]. Under mild regularity conditions, if $\|\alpha_n - \alpha_0\| = o_p(1)$, $\|\mu_n - \mu_0\| = o_p(1)$, and $\|\alpha_n - \alpha_0\| \, \|\mu_n - \mu_0\| = o_p(n^{-1/2})$, then τ_n is asymptotically linear for τ_0 with influence function D_0, i.e., $\tau_n - \tau_0 = P_n D_0 + o_p(n^{-1/2})$. Hence, by the Central Limit Theorem, τ_n is root-n asymptotically normal, and, by Theorem 2.1, it is semiparametrically efficient. The estimator τ_n is additionally doubly robust for consistency: it converges in probability to τ_0 provided that either $\|\alpha_n - \alpha_0\| = o_p(1)$ or $\|\mu_n - \mu_0\| = o_p(1)$. The required regularity conditions amount to negligibility of certain empirical-process remainder terms, which can be ensured, for example, by sample splitting or cross-fitting [Schick, 1986, van der Laan and Rose, 2011, Chernozhukov et al., 2018b].

Rate double robustness does *not* yield doubly robust inference: valid Wald-type inference requires that *both* nuisance estimators be consistent and satisfy the product-rate condition

above (e.g., if each converges at $o_p(n^{-1/4})$). By contrast, under one-sided correctness (i.e., when only one nuisance estimator is consistent, or when one converges too slowly), τ_n remains consistent but the product-rate condition may fail. In this case, τ_n can exhibit non-negligible asymptotic bias that precludes asymptotic linearity, invalidating Wald-type confidence intervals and typically causing undercoverage. This gap motivates the stronger objective of *doubly robust asymptotic linearity*: estimators that remain asymptotically linear whenever *either* nuisance estimator is sufficiently accurate, allowing the other nuisance estimator to be inconsistent or to converge arbitrarily slowly.

2.3 Preliminaries on doubly robust inference

2.3.1 Statistical objective: doubly robust asymptotic linearity

We aim to construct estimators of τ_0 that are doubly robust asymptotically linear (DRAL), meaning they remain asymptotically linear whenever either

(i) $\|\alpha_n - \alpha_0\| = o_p(n^{-1/4})$ and μ_n converges in probability to some $\overline{\mu}_0$, or

(ii) $\|\mu_n - \mu_0\| = o_p(n^{-1/4})$ and α_n converges in probability to some $\overline{\alpha}_0$.

The DRAL property strengthens rate double robustness by allowing one nuisance estimator to converge arbitrarily slowly, or even to be inconsistent. In particular, it yields *best-rate double robustness*: it replaces the usual product-rate condition $\|\alpha_n - \alpha_0\| \cdot \|\mu_n - \mu_0\| = o_p(n^{-1/2})$ with the one-sided requirement $\|\alpha_n - \alpha_0\|^2 \wedge \|\mu_n - \mu_0\|^2 = o_p(n^{-1/2})$. By contrast, the product-rate condition couples the convergence rates of the two nuisances: if one nuisance estimator converges more slowly than $n^{-1/4}$, then the other must converge correspondingly faster than $n^{-1/4}$ to compensate. Without additional assumptions on the data-generating distribution, rate double robustness cannot be improved upon [Balakrishnan et al., 2023, Jin and Syrgkanis, 2024, 2025]. We therefore seek estimators that are DRAL under additional assumptions, while retaining rate double robustness when these conditions fail.

To illustrate the improvement over rate double robustness, suppose that μ_n and α_n converge to μ_0 and α_0 at rates $n^{-\beta/(2\beta+d)}$ and $n^{-\gamma/(2\gamma+d)}$, where d is the dimension of W and $\beta, \gamma > 0$ are smoothness exponents for the nuisance functions. Such rates are

attainable, for example, with local polynomial and random-forest-type methods when μ_0 and α_0 are Hölder smooth [Stone, 1982, Wainwright, 2019, Biau, 2012, Mourtada et al., 2020]. Figure 2.2 visualizes the qualitative relationship between β and γ needed to achieve asymptotic linearity under rate double robustness and DRAL. To satisfy the product-rate condition underlying rate double robustness, one typically requires $\beta + \gamma > d$ [Robins et al., 2008]; thus, if one nuisance is rough ($\beta \wedge \gamma \ll d/2$), the other must be very smooth to compensate. By contrast, DRAL requires only $\beta \vee \gamma > d/2$, so asymptotic linearity can hold even when one nuisance is very rough, provided the other is sufficiently smooth. An analogous separation holds under sparsity: if α_0 and μ_0 have sparsities s_α and s_μ in ambient dimension p, then rate double robustness typically requires $\sqrt{s_\alpha s_\mu} \log p = o(\sqrt{n})$, whereas DRAL requires only $(s_\alpha \wedge s_\mu) \log p = o(\sqrt{n})$ [Bradic et al., 2019a, Smucler et al., 2019, Liu et al., 2023]. More generally, these examples show how the estimator-based DRAL condition can be translated, through known nuisance convergence guarantees, into sufficient conditions on nuisance model classes; in this rate-based sense, DRAL provides an infinite-dimensional model double robustness property.

To illustrate the improvement over rate double robustness, suppose that μ_n and α_n converge to μ_0 and α_0 at rates $n^{-\beta/(2\beta+d)}$ and $n^{-\gamma/(2\gamma+d)}$, where d is the dimension of W and $\beta, \gamma > 0$ are smoothness exponents for the nuisance functions. Such rates are attainable, for example, with local polynomial and random-forest-type methods when μ_0 and α_0 are Hölder smooth [Stone, 1982, Wainwright, 2019, Biau, 2012, Mourtada et al., 2020]. Figure 2.2 visualizes the qualitative relationship between β and γ needed for asymptotic linearity under rate double robustness and DRAL. Under the usual product-rate condition, one typically requires $\beta + \gamma > d$ [Robins et al., 2008]. Thus, if one nuisance is rough ($\beta \wedge \gamma \ll d/2$), the other must be very smooth to compensate. By contrast, DRAL requires only $\beta \vee \gamma > d/2$, so asymptotic linearity can hold whenever either nuisance lies in a sufficiently smooth nonparametric class, even if the other nuisance is much rougher. This gives DRAL a model-robust interpretation at the level of sufficient conditions: one needs one well-estimated nuisance component, rather than a joint smoothness tradeoff between both components. An analogous separation holds under sparsity: if α_0 and μ_0 have sparsities s_α and s_μ in ambient dimension p, then rate double robustness typically requires $\sqrt{s_\alpha s_\mu} \log p = o(\sqrt{n})$, whereas

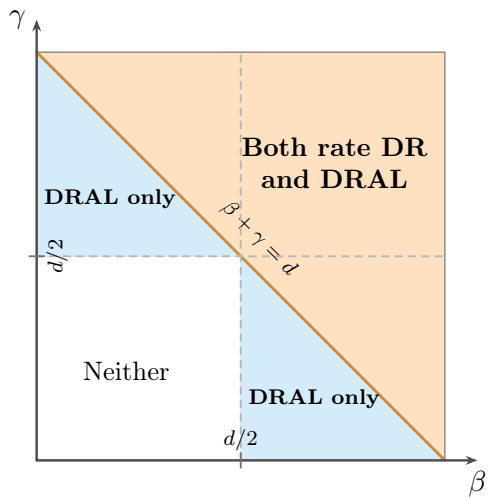


Figure 2.2: Smoothness trade-offs for asymptotic linearity. Under Hölder-type rates $n^{-\beta/(2\beta+d)}$ and $n^{-\gamma/(2\gamma+d)}$, rate double robustness typically requires $\beta+\gamma > d$, whereas DRAL requires only $\beta \vee \gamma > d/2$.

DRAL requires only $(s_\alpha \wedge s_\mu)\log p = o(\sqrt{n})$ [Bradic et al., 2019a, Smucler et al., 2019, Liu et al., 2023].

To elucidate how doubly robust asymptotic linearity can arise, we study the one-step debiased estimator τ_n under possible nuisance misspecification. Suppose that $\alpha_n \to_p \overline{\alpha}_0$ and $\mu_n \to_p \overline{\mu}_0$ in the sense that $\|\alpha_n - \overline{\alpha}_0\| = o_p(1)$ and $\|\mu_n - \overline{\mu}_0\| = o_p(1)$, where at least one limit is correct, i.e., either $\overline{\mu}_0 = \mu_0$ or $\overline{\alpha}_0 = \alpha_0$. Under sample splitting, or under suitable estimator-complexity conditions, we obtain a von Mises-type expansion [Benkeser et al., 2017]

$$\tau_n - \tau_0 = (P_n - P_0)\,\overline{D}_0 + o_p(n^{-1/2}) + \langle \alpha_0 - \alpha_n, \mu_n - \mu_0 \rangle, \tag{2.3}$$

where $\overline{D}_0 := \phi_{0,\overline{\mu}_0} + D_{\overline{\mu}_0,\overline{\alpha}_0}$. Thus, up to the $o_p(n^{-1/2})$ term, the only obstruction to asymptotic linearity of τ_n is the cross-product remainder $\langle \alpha_0 - \alpha_n, \mu_n - \mu_0 \rangle$. Most of the existing literature controls this remainder via rate double robustness, imposing the product-rate condition $\|\alpha_n - \alpha_0\|\,\|\mu_n - \mu_0\| = o_p(n^{-1/2})$, which typically requires both nuisance estimators to be consistent and to converge sufficiently fast [Robins et al., 2000, 1995, van der Laan and Rose, 2011, Chernozhukov et al., 2018a]. If both limits are correct, so that $\overline{\mu}_0 = \mu_0$ and $\overline{\alpha}_0 = \alpha_0$, then $\overline{D}_0$ coincides with the efficient influence function D_0 in Theorem 2.1, yielding efficiency.

When the product-rate condition fails—because one nuisance estimator is inconsistent or

converges too slowly—the cross-product remainder in (2.3) is no longer negligible and can dominate $n^{1/2}(\tau_n - \tau_0)$, precluding asymptotic linearity of the one-step estimator. Achieving DRAL therefore requires additional control or debiasing of $\langle \alpha_0 - \alpha_n, \mu_n - \mu_0 \rangle$ so that it is negligible, or admits an asymptotically linear representation, even without $\|\alpha_n - \alpha_0\| \, \|\mu_n - \mu_0\| = o_p(n^{-1/2})$. In the remainder of this section, we analyze this term and show how additional debiasing yields DRAL whenever either (i) or (ii) holds, under suitable conditions. Readers primarily interested in the estimator may wish to proceed directly to Section 2.4.

2.3.2 General strategy: debiasing the cross-product remainder

A learner-agnostic bias expansion under calibration and partial orthogonality

We now describe a general strategy for debiasing the cross-product remainder $\langle \alpha_0 - \alpha_n, \mu_n - \mu_0 \rangle$ by calibrating the nuisance estimators α_n and μ_n. The goal is to construct these estimators so that the cross-product remainder admits an asymptotically linear representation, behaving like the empirical mean of a mean-zero, possibly degenerate, influence function. This strategy builds on van der Laan [2014] and subsequent work [Benkeser et al., 2017, Díaz and van der Laan, 2017, Díaz, 2019, Dukes et al., 2024], which construct DRAL estimators of average treatment effects when one nuisance estimator may be inconsistent. We extend this line of work in two directions: we consider a broad class of statistical parameters beyond the average treatment effect, and we analyze regimes in which both nuisance estimators are consistent but one may converge arbitrarily slowly. In these settings, we establish asymptotic linearity of the cross-product remainder and give refined conditions under which the remaining higher-order terms are negligible.

We use the following notation throughout this subsection. Denote the nuisance errors by

$$\Delta_{\alpha,n} := \alpha_n - \alpha_0, \qquad \Delta_{\mu,n} := \mu_n - \mu_0.$$

For any map $v : \mathcal{A} \times \mathcal{W} \to \mathbb{R}^k$, let Π_v denote the $L^2(P_0)$-projection onto $\{h \circ v : h : \mathbb{R}^k \to \mathbb{R}\}$. When v and f are nonrandom,

$$(\Pi_v f)(a, w) = E_0\{f(A, W) \mid v(A, W) = v(a, w)\}.$$

Define the projected Riesz and regression errors by

$$s_{n,0} := \Pi_{\mu_n}(\alpha_0 - \alpha_n) = -\Pi_{\mu_n}\Delta_{\alpha,n}, \qquad r_{n,0} := \Pi_{\alpha_n}(\mu_0 - \mu_n) = -\Pi_{\alpha_n}\Delta_{\mu,n}.$$

Thus, $s_{n,0}$ is the component of the Riesz error $\alpha_0 - \alpha_n$ explained by $\mu_n(A, W)$, while $r_{n,0}$ is the component of the regression error $\mu_0 - \mu_n$ explained by $\alpha_n(A, W)$.

Our debiasing strategy calibrates the nuisance estimators so that their errors are empirically orthogonal along the projected directions $s_{n,0}$ and $r_{n,0}$. Specifically, we construct calibrated estimators μ_n and α_n satisfying

$$\frac{1}{n}\sum_{i=1}^{n} s_{n,0}(A_i, W_i)\{Y_i - \mu_n(A_i, W_i)\} = 0, \tag{2.4}$$

$$\psi_n(r_{n,0}) - \frac{1}{n}\sum_{i=1}^{n} r_{n,0}(A_i, W_i)\alpha_n(A_i, W_i) = 0. \tag{2.5}$$

The first condition orthogonalizes the regression residuals against the projected Riesz error $s_{n,0}$. The second enforces the empirical Riesz representation $\psi_n(r_{n,0}) = \langle \alpha_n, r_{n,0}\rangle_{P_n}$ along the projected regression error $r_{n,0}$. If $\psi_n(\cdot) = \langle \cdot, \alpha_n^\star\rangle_{P_n}$ for an empirical representer $\alpha_n^\star$, then the second condition is equivalently $\langle r_{n,0}, \alpha_n^\star - \alpha_n\rangle_{P_n} = 0$.

For the average treatment effect, Benkeser et al. [2017] approximately satisfied analogous conditions using parameter-specific TMLE-style procedures based on estimates of $s_{n,0}$ and $r_{n,0}$; see Appendix 2.D.6. In the next section, we show that isotonic calibration provides a simple way to enforce these conditions. For now, we treat μ_n and α_n satisfying (2.4)–(2.5) as given.

To build intuition, we first study the cross-product remainder under (2.4) and show how this condition reduces sensitivity to estimation error in α_n when μ_n is consistent for μ_0. Following van der Laan [2014] and related work, we begin with the observation that the cross-product remainder depends only on the component of the Riesz estimation error $\alpha_n(A, W) - \alpha_0(A, W)$ lying in the linear span of $\mu_n(A, W)$ and $\mu_0(A, W)$. Specifically, by the law of iterated expectations, conditioning on $(\mu_n(A, W), \mu_0(A, W))$ under P_0 yields

$$\langle \alpha_0 - \alpha_n, \mu_n - \mu_0\rangle = \langle \Pi_{(\mu_n,\mu_0)}(\alpha_0 - \alpha_n), \mu_n - \mu_0\rangle.$$

The orthogonality condition (2.4) then applies a minimal correction to the identifiable component $\Pi_{\mu_n}(\alpha_0 - \alpha_n)$ of the projected error $\Pi_{(\mu_n,\mu_0)}(\alpha_0 - \alpha_n)$, that is, the part determined by $\mu_n(A, W)$. Specifically, following Benkeser et al. [2017], Dukes et al. [2024], and Bonvini

et al. [2024], we decompose the preceding display as

$$\begin{aligned}\langle \alpha_0 - \alpha_n, \mu_n - \mu_0 \rangle &= \langle \Pi_{\mu_n}(\alpha_0 - \alpha_n), \mu_n - \mu_0 \rangle + \langle (\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})(\alpha_0 - \alpha_n), \mu_n - \mu_0 \rangle \\ &= \langle s_{n,0}, \mu_n - \mu_0 \rangle + \langle (\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})\Delta_{\alpha,n}, \mu_0 - \Pi_{\mu_n}\mu_0 \rangle,\end{aligned} \tag{2.6}$$

where the second term is negligible when μ_0 provides little information about the residual $\Delta_{\alpha,n}$ beyond what is already captured by μ_n. The first term is the leading bias component and can prevent asymptotic linearity of τ_n when α_0 is estimated too slowly or inconsistently. The key role of (2.4) is to convert this leading term into an empirical process fluctuation:

$$\langle s_{n,0}, \mu_n - \mu_0 \rangle = -P_0 A_{n,0} = (P_n - P_0)A_{n,0} = (P_n - P_0)A_0 + (P_n - P_0)(A_{n,0} - A_0),$$

where,

$$A_{n,0}(z) := \Pi_{\mu_n}(\alpha_0 - \alpha_n)(a,w)\{y - \mu_n(a,w)\}, \qquad A_0(z) := \Pi_{\mu_0}(\alpha_0 - \overline{\alpha}_0)(a,w)\{y - \mu_0(a,w)\}.$$

The second equality uses $P_n A_{n,0} = 0$, as implied by (2.4). Provided that $\|A_{n,0} - A_0\| = o_p(1)$ and $A_{n,0}$ satisfies an appropriate empirical process condition, the final term on the right is $o_p(n^{-1/2})$. We verify this condition for isotonic calibration in our proofs; see also Appendix 2.D.2.

The following theorem formalizes the above by deriving a bias expansion showing how (2.4) and (2.5) debias the cross-product remainder. To control the higher-order terms in (2.6) and related expansions, we impose the following high-level condition, with sufficient conditions given later. Informally, we require that the estimation error of the slower or inconsistent nuisance be asymptotically orthogonal after partialling out the other nuisance:

B1) *Asymptotic partial orthogonality:* At least one of the following holds:

i) $\langle (\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})\Delta_{\alpha,n}\,,\ \mu_0 - \Pi_{\mu_n}\mu_0 \rangle = O_p(\|\mu_n - \mu_0\|^2)$,

ii) $\langle (\Pi_{(\alpha_n,\alpha_0)} - \Pi_{\alpha_n})\Delta_{\mu,n}\,,\ \alpha_0 - \Pi_{\alpha_n}\alpha_0 \rangle = O_p(\|\alpha_n - \alpha_0\|^2)$.

The term in B1i is a covariance between the component of $\Pi_{(\mu_n,\mu_0)}\Delta_{\alpha,n}$ that remains after partialling out functions of μ_n and the approximation error $\mu_0 - \Pi_{\mu_n}\mu_0$. When this condition holds, the higher-order term in (2.6) is $O_p(\|\mu_n - \mu_0\|^2)$, and is therefore asymptotically negligible when $\|\mu_n - \mu_0\| = o_P(n^{-1/4})$, regardless of the quality of α_n. Similarly, Condition B1ii ensures a related higher-order term is negligible with the roles of α and μ reversed. We

provide sufficient conditions in Section 2.3.2. As a concrete example, Condition B1i holds if $\Pi_{(\mu_n,\mu_0)}\Delta_{\alpha,n}$ admits the approximate partially linear representation

$$\Pi_{(\mu_n,\mu_0)}\Delta_{\alpha,n} = g_n \circ \mu_n + \beta_n \circ (\mu_n, \mu_0) \cdot (\mu_n - \mu_0) + o_p(\|\mu_n - \mu_0\|) \tag{2.7}$$

in $L^2(P_0)$, for some $g_n \colon \mathbb{R} \to \mathbb{R}$ and $\beta_n \colon \mathbb{R}^2 \to \mathbb{R}$ satisfying $\|\beta_n\|_\infty = O_p(1)$. In particular, this condition holds if $(\Pi_{(\mu_n,\mu_0)}\Delta_{\alpha,n})(A, W)$ is sufficiently smooth in $\mu_n(A, W)$ and $\mu_0(A, W)$.

The following theorem involves two additional nuisance functions:

$$\begin{aligned} s_0(a,w) &:= E_0\{\alpha_0(A,W) - \overline{\alpha}_0(A,W) \mid \mu_0(A,W) = \mu_0(a,w)\}, \\ r_0(a,w) &:= E_0\{\mu_0(A,W) - \overline{\mu}_0(A,W) \mid \alpha_0(A,W) = \alpha_0(a,w)\}. \end{aligned} \tag{2.8}$$

Equivalently, $s_0 = \Pi_{\mu_0}(\alpha_0 - \overline{\alpha}_0)$ and $r_0 = \Pi_{\alpha_0}(\mu_0 - \overline{\mu}_0)$. Define

$$B_{n,0}(z) := \phi_{r_{n,0}}(z) + \psi_0(r_{n,0}) - r_{n,0}(a,w)\alpha_n(a,w), \qquad B_0(z) := \phi_{r_0}(z) + \psi_0(r_0) - r_0(a,w)\alpha_0(a,w).$$

The theorem does not require μ_n and α_n to converge to $\overline{\mu}_0$ and $\overline{\alpha}_0$, but we impose this convergence in all subsequent results and discussion.

Theorem 2.3 (Expansion of cross-product remainder under orthogonality)**.** *Both of the following are valid:*

(i) Regression-based decomposition: *If μ_n satisfies (2.4),*

$$\langle \alpha_0 - \alpha_n, \mu_n - \mu_0 \rangle = P_n A_0 + (P_n - P_0)(A_{n,0} - A_0) + \langle (\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})\Delta_{\alpha,n},\, \mu_0 - \Pi_{\mu_n}\mu_0 \rangle$$

If B1i holds, then $\langle (\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})\Delta_{\alpha,n},\, \mu_0 - \Pi_{\mu_n}\mu_0 \rangle = O_p\left(\|\mu_n - \mu_0\|^2 \wedge \|\Pi_{\mu_n}\mu_0 - \mu_0\|\|\alpha_n - \alpha_0\|\right)$.

(ii) Riesz-representer-based decomposition: *If α_n satisfies (2.5),*

$$\langle \alpha_0 - \alpha_n, \mu_n - \mu_0 \rangle = P_n B_0 + (P_n - P_0)(B_{n,0} - B_0) + \langle (\Pi_{(\alpha_n,\alpha_0)} - \Pi_{\alpha_n})\Delta_{\mu,n},\, \alpha_0 - \Pi_{\alpha_n}\alpha_0 \rangle + Rem_{\psi_n}(r_{n,0}),$$

with $Rem_{\psi_n}(r_{n,0}) := \psi_n(r_{n,0}) - \psi_0(r_{n,0}) - P_n\phi_{r_{n,0}}$. *If B1ii holds, then* $\langle (\Pi_{(\alpha_n,\alpha_0)} - \Pi_{\alpha_n})\Delta_{\mu,n},\, \Delta_{\alpha,n} \rangle = O_p\left(\|\alpha_n - \alpha_0\|^2 \wedge \|\mu_n - \mu_0\|\|\Pi_{\alpha_n}\alpha_0 - \alpha_0\|\right)$.

When the nuisance estimators satisfy the empirical orthogonality conditions in (2.4) and (2.5), Theorem 2.3 decomposes the cross-product remainder into an asymptotically linear

term, an empirical-process remainder, and a higher-order cross-product remainder. Under the regression-based decomposition, these terms are $P_n A_0$, $(P_n - P_0)(A_{n,0} - A_0)$, and $\langle(\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})\Delta_{\alpha,n},\, \mu_0 - \Pi_{\mu_n}\mu_0\rangle$. Recalling (2.3), if the remainder terms are $o_p(n^{-1/2})$, then the one-step debiased estimator τ_n, constructed from the orthogonalized nuisance estimators μ_n and α_n, is DRAL with influence function $\overline{D}_0 + 1(\overline{\mu}_0 \neq \mu_0)\, A_0 + 1(\overline{\alpha}_0 \neq \alpha_0)\, B_0$. Moreover, if both nuisance estimators are consistent, so that $\overline{\alpha}_0 = \alpha_0$ and $\overline{\mu}_0 = \mu_0$, then $A_0 = 0$ and $B_0 = 0$. Hence, τ_n is asymptotically linear with influence function $\overline{D}_0$, provided the cross-product remainder is purely second order.

The higher-order cross-product remainder can be bounded using the Cauchy–Schwarz inequality. This is the same approach typically used to bound the traditional cross-product remainder, for which it yields $|\langle \alpha_0 - \alpha_n, \mu_n - \mu_0\rangle| \leq \|\alpha_n - \alpha_0\|\, \|\mu_n - \mu_0\|$ [van der Laan and Rose, 2011, Chernozhukov et al., 2018b, Bonvini et al., 2024]. For the higher-order cross-product remainder, Cauchy–Schwarz yields

$$\left|\langle(\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})\Delta_{\alpha,n},\, \mu_0 - \Pi_{\mu_n}\mu_0\rangle\right| \leq \|(\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})(\alpha_n - \alpha_0)\|\, \|\mu_0 - \Pi_{\mu_n}\mu_0\|. \tag{2.9}$$

This bound is tighter than the bound on $|\langle \alpha_0 - \alpha_n, \mu_n - \mu_0\rangle|$ because it projects out the components of the nuisance errors explained by μ_n. Under B1i, this dominance is even clearer, since then (2.9) converges to 0 at least as quickly as the faster of $\|\mu_n - \mu_0\|^2$ and $\|\alpha_n - \alpha_0\|\|\Pi_{\mu_n}\mu_0 - \mu_0\| = O_p(\|\alpha_n - \alpha_0\|\|\mu_n - \mu_0\|)$. Analogous bounds hold for the Riesz-based decomposition, in which α_n is projected out from both errors.

Sufficient conditions for partial orthogonality

A sufficient condition for B1 is that the projected Riesz error $\Pi_{(\mu_n,\mu_0)}\Delta_{\alpha,n}$ admits the partially linear expansion in (2.7). At a high level, B1i rules out near-perfect alignment between the projected-out regression error and the Riesz error. For example, if $\mu_0 - \Pi_{\mu_n}\mu_0 = \varepsilon_n \Delta_{\alpha,n}/\|\Delta_{\alpha,n}\|$ for some $\varepsilon_n = o(1)$, then $\langle(\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})\Delta_{\alpha,n}, \mu_0 - \Pi_{\mu_n}\mu_0\rangle$ is of order $\|\Delta_{\alpha,n}\|\, \|\mu_0 - \Pi_{\mu_n}\mu_0\|$, which need not be $o_p(\|\mu_n - \mu_0\|^2)$, especially when $\|\Delta_{\alpha,n}\|$ does not converge to zero in probability. Conditions excluding this type of behavior already appear in related work; see, for example, Condition (v) in Appendix F of Benkeser et al. [2020], Theorem 1 of Benkeser et al. [2017], Assumption 3 of Dukes et al. [2024], Lemma I.1 of

Wang et al. [2023], and Section 3.1 of Bonvini et al. [2024]. We next adapt these ideas to our setting and present sufficient conditions in increasing order of strength.

B1* *Projection error coupling* [van der Laan, 2014]:

i) $\|(\Pi_{\mu_n,\mu_0} - \Pi_{\mu_n})\Delta_{\alpha,n}\| = O_p(\|\mu_n - \mu_0\|)$;

ii) $\|(\Pi_{\alpha_n,\alpha_0} - \Pi_{\alpha_n})\Delta_{\mu,n}\| = O_p(\|\alpha_n - \alpha_0\|)$.

The above condition is also assumed in Benkeser et al. [2017, Theorem 1] and Dukes et al. [2024, Assumption 3]; related conditions appear in Díaz and van der Laan [2017] and Díaz [2019].

B1** *Lipschitz continuity* [Bonvini et al., 2024]: Denoting $\mathcal{D}_n := \{Z_1, \ldots, Z_n\}$, it holds that, for all n large enough:

i) $(\widehat{m}, m) \mapsto E_0\{\Delta_{\alpha,n}(A,W) \,|\, \mu_n(A,W) = \widehat{m}, \mu_0(A,W) = m, \mathcal{D}_n\}$ is P_0–almost surely Lipschitz continuous with uniformly bounded constant;

ii) $(\widehat{a}, a) \mapsto E_0\{\Delta_{\mu,n}(A,W) \,|\, \alpha_n(A,W) = \widehat{a}, \alpha_0(A,W) = a, \mathcal{D}_n\}$ is P_0–almost surely Lipschitz continuous with uniformly bounded constant.

The above condition is also used in Bonvini et al. [2024, Section 3.1] and related bounded derivative conditions were also imposed in Benkeser et al. [2020, Condition v in Appendix F], van der Laan et al. [2023, Lemma 1], and Wang et al. [2023, Lemma I.1] to analyze regressions involving estimated features.

Lemma 2.4 (Sufficient conditions for B1)**.** *Conditions B1i and B1ii are implied by B1***i and B1***ii, respectively. Moreover, B1***i and B1***ii are implied by B1****i and B1****ii, respectively.*

Roughly, Condition B1*i requires that $\Pi_{(\mu_n,\mu_0)}\Delta_{\alpha,n}$ be well approximated by $\Pi_{\mu_n}\Delta_{\alpha,n}$ whenever μ_n converges to μ_0 in norm. Heuristically, it asserts that, as $\|\mu_n - \mu_0\| \to 0$, the information in $\{\mu_n(A,W), \mu_0(A,W), Z_1, \ldots, Z_n\}$ for predicting $\Delta_{\alpha,n}$ is asymptotically equivalent to that in $\{\mu_n(A,W), Z_1, \ldots, Z_n\}$. While such equivalences need not hold in general, a sufficient condition is Condition B1**, which ensures that $\Pi_{(\mu_n,\mu_0)}\Delta_{\alpha,n}$

and $\Pi_{(\alpha_n,\alpha_0)}\Delta_{\mu,n}$ are almost surely Lipschitz continuous as bivariate transformations of (μ_n, μ_0) and (α_n, α_0), respectively. By symmetry between μ_n and μ_0, Condition B1**i also implies $\|(\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_0})\Delta_{\alpha,n}\| = O_p(\|\mu_n - \mu_0\|)$ and hence, by the triangle inequality, $\|(\Pi_{\mu_0} - \Pi_{\mu_n})\Delta_{\alpha,n}\| = O_p(\|\mu_n - \mu_0\|)$. An analogous conclusion holds for Condition B1**ii. Conditions B1–B1** can also be relaxed to Hölder-type continuity [Bonvini et al., 2024]; in that case, a weaker form of the bias expansion in Theorem 2.3, governed by the Hölder exponent, can still be established.

More broadly, some form of structural assumption is unavoidable for meaningfully improving upon standard doubly robust estimators, since first-order methods are optimal in structure-agnostic settings [Balakrishnan et al., 2023, Jin and Syrgkanis, 2025]. Conditions B1–B1** provide one such set of assumptions. For comparison, analyses of higher-order influence function estimators [Robins et al., 2008] typically impose Hölder-type smoothness assumptions on the potentially high-dimensional error functions $\Delta_{\mu,n}$ and $\Delta_{\alpha,n}$, whereas sparsity-based methods require approximate sparsity and linearity of the nuisance models [Liu et al., 2023]. Our guarantees instead rely on the *asymptotic partial orthogonality* condition in B1, which is qualitatively different from the smoothness and approximability assumptions used in these approaches. In particular, the sufficient Condition B1** requires only smoothness of certain bivariate transformations of the true and estimated nuisances. Informally, Condition B1**i may be plausible when the joint level sets $\{(a, w) : \mu_n(a, w) = \widehat{m},\ \mu_0(a, w) = m\}$ vary smoothly with $(\widehat{m}, m)$, and the conditional mean of $\Delta_{\alpha,n}$ over these level sets depends continuously on $(\widehat{m}, m)$. Finally, (2.9) shows that, even when these additional conditions do not hold, the orthogonality conditions typically still attenuate the bias of the one-step debiased estimator.

2.4 Calibrated debiased machine learning

2.4.1 Enforcing orthogonality via calibration

We now establish a key insight of this paper: the empirical orthogonality equations (2.4) and (2.5) can be enforced by calibrating the nuisance estimators μ_n and α_n. Combined with Theorem 2.3 of the previous subsection, this reformulation shows that calibration enables

the construction of DRAL debiased machine learning estimators, provided the nuisance estimators satisfy the partial orthogonality conditions in B1. In particular, calibration implicitly performs the bias correction needed for the error decomposition in Theorem 2.3 to apply. We now formalize this connection and introduce our calibrated debiased machine learning framework.

Calibration is a widely used post-hoc technique in predictive modeling that transforms a model's fitted values to better align with the target they are intended to predict [Lichtenstein et al., 1977, Zadrozny and Elkan, 2002, Niculescu-Mizil and Caruana, 2005b, Bella et al., 2010, Guo et al., 2017, Gupta et al., 2020]. Motivated by this idea, we say that μ_n and α_n are *(perfectly) empirically calibrated* with respect to the least-squares and Riesz losses if

$$\sum_{i=1}^{n}\{Y_i - \mu_n(A_i, W_i)\}^2 = \min_{\theta}\sum_{i=1}^{n}\{Y_i - \theta(\mu_n(A_i, W_i))\}^2, \tag{2.10}$$

$$\sum_{i=1}^{n}\{\alpha_n(A_i, W_i)\}^2 - 2\psi_n(\alpha_n) = \min_{\theta}\left\{\sum_{i=1}^{n}\{\theta(\alpha_n(A_i, W_i))\}^2 - 2\psi_n(\theta \circ \alpha_n)\right\}, \tag{2.11}$$

where the minima are taken over all one-dimensional functions $\theta : \mathbb{R} \to \mathbb{R}$. Equivalently, the empirical risk of each nuisance estimator cannot be reduced by post-composing it with an arbitrary one-dimensional mapping [Whitehouse et al., 2024]. Beyond enabling doubly robust inference, empirical calibration can improve estimator stability and predictive accuracy. Calibration for the outcome regression ensures that $\mu_n(a, w)$ equals the empirical mean outcome among individuals with the same predicted value, whereas calibration for the propensity score enforces covariate balance across strata of the estimated propensity score [Deshpande and Kuleshov, 2023, van der Laan et al., 2024b].

Empirical calibration of the nuisance estimators implies the orthogonality conditions (2.4) and (2.5). To ensure that this implication holds for α_n, we assume that the estimated functional ψ_n is linear, as holds automatically for $\psi_n(\mu) := \frac{1}{n}\sum_{i=1}^{n} m(Z_i, \mu)$ in Example 2.2.

B2) *(Linearity of the estimated functional)* P_0–almost surely, $\psi_n : \mathcal{H} \to \mathbb{R}$ satisfies $\psi_n(c\mu_1 + \mu_2) = c\,\psi_n(\mu_1) + \psi_n(\mu_2)$ for each $\mu_1, \mu_2 \in \mathcal{H}$ and $c \in \mathbb{R}$.

Under B2, the first-order optimality conditions for (2.10) and (2.11) yield the equivalent moment conditions: for each transformation $\theta : \mathbb{R} \to \mathbb{R}$,

$$\frac{1}{n}\sum_{i=1}^{n}\theta\left(\mu_n(A_i, W_i)\right)\left\{Y_i - \mu_n(A_i, W_i)\right\} = 0, \tag{2.12}$$

$$\frac{1}{n}\sum_{i=1}^{n}\theta\left(\alpha_n(A_i,W_i)\right)\alpha_n(A_i,W_i)-\psi_n(\theta\circ\alpha_n)=0. \tag{2.13}$$

That is, calibration requires the residuals of μ_n and α_n to be orthogonal to *every* one-dimensional transformation of the corresponding fitted values. In particular, they are orthogonal to the specific transformations $s_{n,0} = \Pi_{\mu_n}(\alpha_0 - \alpha_n)$ and $r_{n,0} = \Pi_{\alpha_n}(\mu_0 - \mu_n)$ appearing in the orthogonality conditions. We summarize this implication in the following lemma.

Lemma 2.5 (Empirical calibration debiases the cross-product remainder)**.** *If μ_n is empirically calibrated so that (2.10) holds, then μ_n satisfies (2.4). If, in addition, B2 holds and α_n satisfies (2.11), then α_n satisfies (2.5).*

We refer to debiased machine learning with calibrated nuisance estimators as *calibrated DML*.

In practice, calibration is implemented by applying a post-hoc one-dimensional map to an initial nuisance estimator, chosen to reduce the empirical loss over a flexible class [Whitehouse et al., 2024]. A simple option is histogram (binning) calibration: partition the fitted values and assign a constant calibrated prediction within each bin, chosen to minimize the loss within that bin (e.g., an empirical mean) [Zadrozny and Elkan, 2001, Gupta and Ramdas, 2021, van der Laan and Alaa, 2025]. A widely used alternative is *isotonic calibration* [Zadrozny and Elkan, 2002, Niculescu-Mizil and Caruana, 2005b], which fits a monotone map and can be viewed as adaptive binning regularized by a monotonicity constraint [Van Der Laan et al., 2023, van der Laan and Alaa, 2025]. It is tuning-free [Guntuboyina and Sen, 2018] and, unlike histogram binning, can only improve or leave unchanged the empirical risk, since the identity map is monotone. Moreover, the calibration properties in (2.10) and (2.11) follow from a basic property of isotonic regression: it yields a piecewise-constant fit, equivalently a histogram-binning estimator on a data-dependent partition of the input values produced by the pooled adjacent violators algorithm [Barlow and Brunk, 1972].

2.4.2 Proposed estimator using isotonic calibration

We now introduce *isocalibrated DML*, a concrete instantiation of calibrated DML that enforces (2.10) and (2.11) exactly through isotonic calibration. The procedure is automatic: it

requires no knowledge of the specific form of the linear functional ψ_0, only black-box access to an asymptotically linear estimator $\psi_n(\mu)$ of $\psi_0(\mu)$ for any fixed μ.

Our post-hoc procedure uses isotonic regression [Barlow and Brunk, 1972] to calibrate an initial outcome regression μ_n and Riesz representer estimator α_n. Specifically, we define the isotonic-calibrated ('isocalibrated') nuisance estimators μ_n^* and α_n^* by

$$\begin{aligned} \mu_n^* &:= f_n \circ \mu_n, \quad \text{where} \quad f_n \in \operatorname*{argmin}_{f \in \mathcal{F}_{\text{iso}}} \sum_{i=1}^{n} \{Y_i - f(\mu_n(A_i, W_i))\}^2, \\ \alpha_n^* &:= g_n \circ \alpha_n, \quad \text{where} \quad g_n \in \operatorname*{argmin}_{g \in \mathcal{F}_{\text{iso}}} \sum_{i=1}^{n} \{g(\alpha_n(A_i, W_i))\}^2 - 2\psi_n(g \circ \alpha_n), \end{aligned} \tag{2.14}$$

where $\mathcal{F}_{\text{iso}}$ denotes the class of real-valued monotone nondecreasing functions on $\mathbb{R}$. Our proposed isocalibrated DML estimator is then

$$\tau_n^* := \psi_n(\mu_n^*) + \frac{1}{n} \sum_{i=1}^{n} \alpha_n^*(A_i, W_i)\{Y_i - \mu_n^*(A_i, W_i)\}, \tag{2.15}$$

that is, the debiased machine learning estimator formed using the calibrated nuisance estimators.

To obtain theoretical guarantees without empirical process conditions, we use cross-fitting to construct the initial nuisance estimators [van der Laan et al., 2011, Chernozhukov et al., 2018a]. Algorithm 1 displays the cross-fitted version of (2.15); the only change is that cross-fitted initial nuisance estimates are used for calibration and to compute τ_n^*. The calibration step is still performed on the full sample so that the orthogonality conditions hold in sample. This does not undermine the cross-fitting guarantees: monotone function classes have finite uniform entropy integrals, which we use to control empirical process terms arising in our analysis [Rabenseifner et al., 2025]. Cross-fitting can produce nuisance estimates that are poorly empirically calibrated on holdout data, and full-sample calibration corrects this.

The isotonic regression solutions f_n and g_n are uniquely defined only at the observed values of μ_n and α_n, and any such solution guarantees empirical calibration. Following Groeneboom and Lopuhaa [1993], we take the unique càdlàg piecewise-constant versions with jumps only at these observed values. Computing g_n in (2.14) may involve a nonstandard loss, but it can be obtained as the empirical risk minimizer over monotone univariate regression trees of unbounded depth. Thus, (2.14) can be solved using regression-tree software that supports monotonicity constraints and custom losses, such as the `R` and `Python`

Algorithm 1 Calibrated DML using isotonic calibration

Input: dataset $\mathcal{D}_n = \{O_i : i = 1, \ldots, n\}$, number J of cross-fitting splits

1: partition $\mathcal{D}_n$ into datasets $\mathcal{C}^{(1)}, \mathcal{C}^{(2)}, \ldots, \mathcal{C}^{(J)}$;
2: **for** $s = 1, \ldots, J$ **do**
3: get initial estimators $\alpha_{n,s}$ of α_0 and $\mu_{n,s}$ of μ_0 from $\mathcal{E}^{(s)} := \mathcal{D}_n \backslash \mathcal{C}^{(s)}$;
4: get initial estimators $\mu \mapsto \psi_{n,s}(\mu)$ of $\mu \mapsto \psi_0(\mu)$ from $\mathcal{D}_n$, where sample-splitting is performed as needed to satisfy Condition C4;
5: **end for**
6: for $s = 1, \ldots, J$, set $j(i) := s$ for each $i \in \mathcal{C}^{(s)}$;
7: compute calibrators f_n and g_n using pooled out-of-fold estimates as

$$f_n \in \operatorname*{argmin}_{f \in \mathcal{F}_{\text{iso}}} \sum_{i=1}^{n} \left\{ Y_i - f(\mu_{n,j(i)}(A_i, W_i)) \right\}^2$$

$$g_n \in \operatorname*{argmin}_{g \in \mathcal{F}_{\text{iso}}} \sum_{i=1}^{n} \left\{ g(\alpha_{n,j(i)}(A_i, W_i))^2 - 2\psi_{n,j(i)}(g \circ \alpha_{n,j(i)}) \right\};$$

8: for $s = 1, \ldots, J$, set $\mu^*_{n,s} := f_n \circ \mu_{n,s}$ and $\alpha^*_{n,s} := g_n \circ \alpha_{n,s}$;
9: compute calibrated one-step debiased estimator

$$\tau^*_n := \frac{1}{J} \sum_{j=1}^{J} \left[\psi_{n,j}(\mu^*_{n,j}) + \frac{1}{n} \sum_{i=1}^{n} \alpha^*_{n,j(i)}(A_i, W_i) \left\{ Y_i - \mu^*_{n,j(i)}(A_i, W_i) \right\} \right]. \tag{2.16}$$

10: **return** τ^*_n

implementations of `xgboost` [Chen and Guestrin, 2016]; see Lee and Schuler [2025] for fitting trees with Riesz losses. Equivalently, g_n can be computed by empirical risk minimization over a basis of threshold indicators with a nonnegative coefficient constraint. The next example shows that, for many causal parameters, this reduces to standard isotonic regression.

Example 2.1 (continued)**.** Consider the G-computation identification $\tau_0(a_0) := E_0\{\mu_0(a_0, W)\}$ of the mean counterfactual outcome Y^{a_0} for $a_0 \in \mathcal{A}$. In this case, the Riesz representer α_0, given by $(a, w) \mapsto 1(a = a_0)/\pi_0(a_0 \mid w)$, is determined by the propensity score π_0. Given estimators μ_n of μ_0 and π_n of π_0, an isocalibrated DML estimator of $\tau_0(a_0)$ (without cross-fitting) is

$$\tau^*_n(a_0) := \frac{1}{n} \sum_{i=1}^{n} \left[\mu^*_n(a_0, W_i) + \alpha^*_n(A_i, W_i)\{Y_i - \mu^*_n(A_i, W_i)\} \right],$$

where, for each $w \in \mathcal{W}$, we set $\mu^*_n(a_0, w) := f_{n,a_0}(\mu_n(a_0, w))$ and $\alpha^*_n(a_0 \mid w) := g_{n,a_0}(\pi_n^{-1}(a_0 \mid w))$ with

$$f_{n,a_0} \in \operatorname*{argmin}_{f \in \mathcal{F}_{\text{iso}}} \sum_{i=1}^{n} 1(A_i = a_0) \left\{ Y_i - f(\mu_n(a_0, W_i)) \right\}^2;$$

$$g_{n,a_0} \in \operatorname*{argmin}_{g \in \mathcal{F}_{\text{iso}}} \sum_{i=1}^{n} \left[\left\{ 1(A_i = a_0)\, g(\pi_n^{-1}(A_i \mid W_i)) \right\}^2 - 2g(\pi_n^{-1}(a_0 \mid W_i)) \right].$$

Here, we optionally stratify the outcome-regression calibration step in (2.14) by treatment, ensuring empirical calibration within each treatment level and hence marginal empirical calibration as well. A useful property of isotonic regression is that, for a broad class of convex losses—namely, Bregman losses—the isotonic fit is *universal*: the solution is invariant to the particular loss used within this class [Jordan et al., 2022a]. Accordingly, the second step can be implemented via logistic balancing-weight regression [Sverdrup and Hastie, 2026].

A similar invariance holds for isotonic calibration of the propensity score, which implicitly also calibrates the corresponding inverse weights, since $x \mapsto 1/x$ is monotone on $(0, \infty)$. In particular, g_{n,a_0} equals $t \mapsto 1/\widetilde{g}_{n,a_0}(t)$, where

$$\widetilde{g}_{n,a_0} \in \operatorname*{argmin}_{g \in \mathcal{F}_{\text{iso}}} \sum_{i=1}^{n} \left\{ 1(A_i = a_0) - g\big(\pi_n(a_0 \mid W_i)\big) \right\}^2 .$$

This equivalence is shown in van der Laan et al. [2024b] using the first-order conditions that characterize isotonic regression solutions g_{n,a_0} and $\widetilde{g}_{n,a_0}$. The calibrated propensities $\{\widetilde{g}_{n,a_0}(\pi_n(A_i \mid W_i))\}_{i=1}^n$ may be zero for observations with $A_i \neq a_0$, and thus induce infinite weights; however, this is not problematic, since the AIPW estimator only requires evaluating the corresponding weights for observations with $A_i = a_0$. `Python` and `R` implementations of this calibration procedure are provided in Appendix 2.H. □

Our estimator is related to the automatic debiased machine learning (autoDML) estimators of Chernozhukov et al. [2018b, 2022a,c], which use tailored loss functions to learn cross-fitted estimators of the Riesz representer and the outcome regression. AutoDML is a first-order debiasing method that uses the Riesz loss to estimate the Riesz representer nuisance, potentially improving finite-sample performance and nuisance-estimator quality. By itself, however, autoDML does not remove the usual product-rate requirement for asymptotic linearity. In contrast, our calibrated DML framework uses these loss functions in a second stage to *calibrate* user-supplied cross-fitted nuisance estimators, yielding DRAL when one nuisance is estimated at a fast enough rate, even if the other is inconsistent. Finally, calibrated DML can be incorporated into the autoDML pipeline by applying isotonic regression to calibrate the cross-fitted nuisance estimates before forming the one-step estimator.

2.4.3 Bootstrap-assisted confidence intervals

In previous works, doubly robust confidence intervals were obtained through a standard Wald construction based on the asymptotic normality of the corrected estimator, using adjusted estimators of the asymptotic variance [Benkeser et al., 2017, Dukes et al., 2024, Bonvini et al., 2024]. Such confidence interval are valid if either nuisance estimator is consistent at a sufficiently fast rate. A drawback of previous approaches is that they require consistently estimating additional nuisance functions not required by the estimator itself. Here, we propose a new method that avoids this requirement. It uses a variant of the bootstrap to construct doubly robust confidence intervals associated with the estimator described in Algorithm 1 [Efron and Tibshirani, 1994].

Besides avoiding estimating new nuisances, our proposed bootstrap-assisted inference procedure is easy to describe and computationally efficient. Holding the initial cross-fitted nuisance estimators fixed, our procedure recomputes the calibrated DML estimator on bootstrap samples, only refitting the isotonic regressions and recomputing the empirical mean used to evaluate the DML estimator. In this way, each bootstrap replicate can be solved in near-linear time using standard software [Barlow and Brunk, 1972]. Consequently, computing our bootstrap procedure is typically much less expensive than fitting the original DML estimator.

We focus our description of the automated algorithm on the case where the linear functional ψ_0 is of the form $\mu \mapsto E_0\{m(Z, \mu)\}$ for a known map $m : \mathcal{Z} \times \mathcal{H} \to \mathbb{R}$. In such cases, we consider the empirical plug-in estimator ψ_n of ψ_0, defined as $\mu \mapsto \frac{1}{n}\sum_{i=1}^{n} m(Z_i, \mu)$. A similar algorithm can be devised for the general case by appropriately bootstrapping the estimation of the functional ψ_n. As in the previous subsection, we first describe our approach in the simpler case where the initial estimators are not cross-fitted. The general procedure allowing for cross-fitting is provided in Algorithm 2. To this end, let α_n and μ_n be initial estimators of α_0 and μ_0 computed once using the original sample. Our approach involves bootstrapping the meta isotonic fitting step of (2.14) used originally to obtain the calibrators f_n and g_n. Specifically, we sample $\{Z_i^{\#} : i = 1, \ldots, n\}$ with replacement from the original sample $\{Z_i : i = 1, \ldots, n\}$ and compute the bootstrapped calibrators

$$f_n^{\#} \in \operatorname*{argmin}_{f \in \mathcal{F}_{\text{iso}}} \sum_{i=1}^{n} \{Y_i^{\#} - f(\mu_n(A_i^{\#}, W_i^{\#}))\}^2;$$

$$g_n^{\#} \in \underset{g \in \mathcal{F}_{\text{iso}}}{\operatorname{argmin}} \sum_{i=1}^{n} \left\{ g(\alpha_n(A_i^{\#}, W_i^{\#}))^2 - 2m(Z_i^{\#}, g \circ \alpha_n) \right\}.$$

Next, we obtain the bootstrap-calibrated nuisance estimators $\alpha_n^{\#} := g_n^{\#} \circ \alpha_n$ and $\mu_n^{\#} := f_n^{\#} \circ \mu_n$, and compute the bootstrapped estimator $\tau_n^{\#}$ of τ_0 from $\mu_n^{\#}$ and $\alpha_n^{\#}$ following the same steps as in (2.15). This procedure is repeated K times to obtain K bootstrap estimators $\tau_{n,1}^{\#}, \ldots, \tau_{n,K}^{\#}$ of τ_0. The choice $K = 10{,}000$ has been recommended for use in practice [Efron and Tibshirani, 1994].

To construct a bootstrap-assisted confidence interval, we first compute the bootstrap average $\overline{\tau}_n^{\#} := \frac{1}{K} \sum_{k=1}^{K} \tau_{n,k}^{\#}$. For given confidence level $\rho \in (0,1)$, we use the interval

$$\text{CI}_n(\rho) := \left(\tau_n^* - q(\rho)\sigma_n^{\#},\ \tau_n^* + q(\rho)\sigma_n^{\#} \right),$$

where $\sigma_n^{\#}$ is the empirical standard deviation of $\tau_{n,1}^{\#}, \ldots, \tau_{n,K}^{\#}$ and, for each $\gamma \in (0,1)$, $q(\gamma)$ denotes the $\gamma/2$–quantile of the standard normal distribution. Alternatively, we could consider a percentile bootstrap interval $(\tau_n^* - q_n(\rho),\ \tau_n^* - q_n(2-\rho))$, where for each $\gamma \in (0,1)$ $q_n(\gamma)$ denotes the empirical $\gamma/2$–quantile of the centered bootstrapped estimates $\tau_{n,1}^{\#} - \overline{\tau}_n^{\#}, \ldots, \tau_{n,K}^{\#} - \overline{\tau}_n^{\#}$; this interval may be more robust to finite sample deviations from normality. We refer to Algorithm 2 in Appendix 2.D.3 for the pseudo-code implementing this approach with cross-fitting.

2.5 Theoretical properties

2.5.1 Notation for cross-fitting and fold-averaging

Let $f_\diamond := \{f_j : j \in [J]\}$ be a collection of fold-specific functions with each $f_j \in L^2(P_0)$, as obtained by cross-fitting. We extend the notation of Section 2.3.2 to fold-specific functions by viewing $f_\diamond : (z, j) \mapsto f_j(z)$ as a single function on the extended domain $\mathcal{Z} \times [J]$. Specifically, we define the fold-averaged operator and the associated L^2 norm by $\overline{P}_0 f_\diamond := J^{-1} \sum_{j=1}^{J} P_0 f_j$ and $\|f_\diamond\|_{\overline{P}_0} := \{\overline{P}_0(f_\diamond^2)\}^{1/2}$. Let $v_\diamond := \{v_j : \mathcal{A} \times \mathcal{W} \to \mathbb{R}^k\}_{j=1}^{J}$ be feature maps, and define the fold-averaged projection operator $\overline{\Pi}_{v_\diamond} : \mathcal{H}^J \to \mathcal{H}^J$ by

$$(\overline{\Pi}_{v_\diamond} f_\diamond)_j := \theta^\star \circ v_j, \qquad \theta^\star \in \arg\min_{\theta} \|f_\diamond - \theta \circ v_\diamond\|_{\overline{P}_0}^2,$$

where the minimization is over measurable $\theta : \mathbb{R}^k \to \mathbb{R}$ such that $\|\theta \circ v_\diamond\|_{\overline{P}_0} < \infty$. Thus, $\overline{\Pi}_{v_\diamond} f_\diamond$ is the $L^2(\overline{P}_0)$ projection of $f_\diamond$ onto functions of $v_\diamond$ using a single map θ shared across

folds.

2.5.2 Doubly robust asymptotic linearity and inference

Main result

We present our main theorem establishing doubly robust asymptotic linearity for the isocalibrated DML estimator in Algorithm 1. Before stating the theorem, we introduce and discuss the required conditions. Let $\mu^*_{n,\diamond} := \{\mu^*_{n,j} : j \in [J]\}$ and $\alpha^*_{n,\diamond} := \{\alpha^*_{n,j} : j \in [J]\}$ denote the isocalibrated nuisance estimators, and define fold-specific errors $\Delta_{\alpha,n,j} := \alpha^*_{n,j} - \alpha_0$ and $\Delta_{\mu,n,j} := \mu^*_{n,j} - \mu_0$ for $j \in [J]$. We define fold-averaged analogues of the covariances in Condition B1 from Section 2.2.2 by

$$\overline{\gamma}_{\mu,0}(\mu^*_{n,\diamond}, \alpha^*_{n,\diamond}) = \left\langle \overline{\Pi}_{(\mu^*_{n,\diamond},\, \mu_0)} \Delta_{\alpha,n,\diamond} - \overline{\Pi}_{\mu^*_{n,\diamond}} \Delta_{\alpha,n,\diamond}\,,\ \mu_0 - \overline{\Pi}_{\mu^*_{n,\diamond}} \mu_0 \right\rangle_{\overline{P}_0},$$
$$\overline{\gamma}_{\alpha,0}(\mu^*_{n,\diamond}, \alpha^*_{n,\diamond}) = \left\langle \overline{\Pi}_{(\alpha^*_{n,\diamond},\, \alpha_0)} \Delta_{\mu,n,\diamond} - \overline{\Pi}_{\alpha^*_{n,\diamond}} \Delta_{\mu,n,\diamond}\,,\ \alpha_0 - \overline{\Pi}_{\alpha^*_{n,\diamond}} \alpha_0 \right\rangle_{\overline{P}_0}.$$

We rely on the following conditions:

C1) *At least one nuisance is well-estimated:* For some $(\overline{\alpha}_0, \overline{\mu}_0) \in \mathcal{H}^2$, either of the following holds:

i) $\|\alpha^*_{n,\diamond} - \alpha_0\|_{\overline{P}_0} = o_p(n^{-1/4})$ and $\|\mu^*_{n,\diamond} - \overline{\mu}_0\|_{\overline{P}_0} = o_p(1)$;

ii) $\|\mu^*_{n,\diamond} - \mu_0\|_{\overline{P}_0} = o_p(n^{-1/4})$ and $\|\alpha^*_{n,\diamond} - \overline{\alpha}_0\|_{\overline{P}_0} = o_p(1)$.

C2) *Partial orthogonality and consistency:* For each $j \in [J]$, assume that the following holds jointly with whichever part of C1 is assumed:

i) Under C1i, $\overline{\gamma}_{\alpha,0}(\mu^*_{n,\diamond}, \alpha^*_{n,\diamond}) = O_p(\|\alpha^*_{n,\diamond} - \alpha_0\|^2_{\overline{P}_0})$ and $\|(\overline{\Pi}_{\alpha^*_{n,\diamond}} - \overline{\Pi}_{\alpha_0})\Delta_{\mu,n,\diamond}\|_{\overline{P}_0} = o_p(1)$.

ii) Under C1ii, $\overline{\gamma}_{\mu,0}(\mu^*_{n,\diamond}, \alpha^*_{n,\diamond}) = O_p(\|\mu^*_{n,\diamond} - \mu_0\|^2_{\overline{P}_0})$ and $\|(\overline{\Pi}_{\mu^*_{n,\diamond}} - \overline{\Pi}_{\mu_0})\Delta_{\alpha,n,\diamond}\|_{\overline{P}_0} = o_p(1)$.

Condition C1 is a DRAL-type rate condition. Under C1i we have $\overline{\alpha}_0 = \alpha_0$, whereas under C1ii we have $\overline{\mu}_0 = \mu_0$. The first part of C2 ensures that a cross-averaged version of the expansion in Theorem 2.3 applies to the cross-fitted calibrated nuisance estimators; sufficient

conditions akin to Lemma 2.4 are given in Appendix A.3.3. The second part is used to control empirical process remainders by ensuring that $r^*_{n,j}$ and $s^*_{n,j}$ are consistent for r_0 and s_0 whenever $\mu^*_{n,j}$ and $\alpha^*_{n,j}$ are, respectively. When both parts of C1 hold, C2 is unnecessary, since the usual product-rate condition applies.

Our results further rely on the following regularity conditions, versions of which are commonly assumed in the debiased machine learning literature [Van Der Laan et al., 2023, Rabenseifner et al., 2025]. We set $r^*_{n,j} := (\overline{\Pi}_{\alpha^*_{n,\diamond}}(\mu_0 - \mu^*_{n,\diamond}))_j$ and $s^*_{n,j} := (\overline{\Pi}_{\mu^*_{n,\diamond}}(\alpha_0 - \alpha^*_{n,\diamond}))_j$ for $j = 1, \ldots, J$, which respectively approximate $r_0 := \Pi_{\alpha_0}(\mu_0 - \overline{\mu}_0)$ and $s_0 := \Pi_{\mu_0}(\alpha_0 - \overline{\alpha}_0)$, defined analogously to (2.8). Let $P_{n,j}$ denote the empirical measure corresponding to the j-th fold $\mathcal{C}^{(j)}$ in the cross-fitting procedure.

C3) *Linearity of estimated functional:* For each $j \in [J]$, the map $\psi_{n,j} : \mathcal{H} \to \mathbb{R}$ is linear P_0-almost surely; that is, $\psi_{n,j}(c\mu_1 + \mu_2) = c\,\psi_{n,j}(\mu_1) + \psi_{n,j}(\mu_2)$ for all $\mu_1, \mu_2 \in \mathcal{H}$ and $c \in \mathbb{R}$.

C4) *Asymptotic linearity of the functional estimator:* For each $j \in [J]$, $\psi_{n,j}(\mu^*_{n,j}) - \psi_0(\mu^*_{n,j}) - P_{n,j}\phi_{0,\overline{\mu}_0} = o_p(n^{-1/2})$ and, under C1i, also $\psi_{n,j}(r^*_{n,j}) - \psi_0(r^*_{n,j}) - P_{n,j}\phi_{0,r_0} = o_p(n^{-1/2})$.

C5) *Boundedness of nuisance functions:* for some $M \in (0, \infty)$, Y and the functions $\mu^*_{n,j}$, $\alpha^*_{n,j}$, μ_0, α_0, $\overline{\mu}_0$ and $\overline{\alpha}_0$ are uniformly bounded by M with P_0–probability tending to one for each $j = 1, \ldots, J$;

C6) *Calibration functions have finite total variation:* There exists $M < \infty$ such that, for all $n \in \mathbb{N}$, there are measurable representatives $\theta_{n,\mu} : \mathbb{R} \to \mathbb{R}$ and $\theta_{n,\alpha} : \mathbb{R} \to \mathbb{R}$ satisfying $\theta_{n,\mu} \circ \mu_{n,\diamond} = \overline{\Pi}_{\mu_{n,\diamond}}\Delta_{\alpha,n,\diamond}$ and $\theta_{n,\alpha} \circ \alpha_{n,\diamond} = \overline{\Pi}_{\alpha_{n,\diamond}}\Delta_{\mu,n,\diamond}$ almost surely, with total variation norms bounded above by M almost surely.

Condition C3 requires the estimated functional to be linear; in most applications, this linearity is inherited from ψ_0 and is therefore satisfied by standard plug-in estimators. Condition C4 is automatic when ψ_0 is known and $\psi_{n,j} = \psi_0$. It also holds, under standard

conditions, for expectation-representable functionals $\psi_0(f) = P_0 m(\cdot, f)$ satisfying (2.2), with known m, when estimated by the empirical plug-in map $\psi_{n,j}(f) = P_{n,j} m(\cdot, f)$; see Lemma A.13 of Appendix A.3.5. The condition is included mainly to allow more general nuisance-dependent linear functionals, for which ψ_0 may itself depend on unknown components of P_0. For example, for a distribution-shift functional $\psi_0(\mu) = E_{P_0^{\text{shift}}}\{m(Z, \mu)\}$, with $P_0^{\text{shift}} \neq P_0$, the condition holds for $\psi_n(f) = P_n^{\text{shift}} m(\cdot, f)$ under similar conditions [Chernozhukov et al., 2023]. When verifying Condition C4 requires additional rate conditions on μ_n, α_n, or other nuisance estimators, the DRAL guarantee applies conditional on those prerequisites.

Conditions C5 and C6 are used to control empirical-process remainders. Condition C5 imposes standard boundedness assumptions and could potentially be relaxed to suitable tail conditions (e.g., sub-Gaussianity). Condition C6 rules out settings in which the optimal $L^2(\overline{P}_0)$ predictors of the estimation errors $\Delta_{\alpha,n,\diamond}$ and $\Delta_{\mu,n,\diamond}$, given only the initial nuisance estimators $\mu_{n,\diamond}$ and $\alpha_{n,\diamond}$, are pathological—in the sense that they have infinite variation norm when viewed as univariate functions of the nuisance estimators [Van Der Laan et al., 2023]. More detailed discussion of these conditions is provided in Appendix 2.B.

We now state our main theorem, which establishes doubly robust asymptotic linearity of the cross-fitted isocalibrated DML estimator. The limiting distribution is governed by the influence function

$$\chi_0 := \phi_{0,\overline{\mu}_0} + D_{\overline{\mu}_0,\overline{\alpha}_0} + A_0 + \Big\{B_0 + \psi_0(\overline{\mu}_0 - \mu_0)\Big\},$$

where $A_0(z) = s_0(a, w)\{y - \mu_0(a, w)\}$ and $B_0(z) = \phi_{0,r_0}(z) + \psi_0(r_0) - r_0(a, w)\alpha_0(w, a)$. When $\psi_0(\mu) = E_0\{m(Z, \mu)\}$, as in (2.2), B_0 simplifies to $m(\cdot, r_0) - r_0\alpha_0$. The first two terms have the same form as the efficient influence function D_0 in Theorem 2.1, with nuisance components replaced by their probability limits. The remaining terms are additional influence-function components induced by nuisance misspecification. Under C1, at most one of these components is nonzero: if $\overline{\alpha}_0 = \alpha_0$, then $A_0 \equiv 0$, whereas if $\overline{\mu}_0 = \mu_0$, then $B_0 + \psi_0(\overline{\mu}_0 - \mu_0) \equiv 0$.

Theorem 2.4 (Main result: DRAL of τ_n^*)**.** *Suppose Conditions A1–A3, C3–C6, and that at least one nuisance is estimated well (C1) with partially orthogonal errors (C2). Then:*

i) The isocalibrated DML estimator τ_n^* of τ_0 is asymptotically linear with influence function χ_0.

ii) If, in addition, $(\overline{\alpha}_0, \overline{\mu}_0) = (\alpha_0, \mu_0)$, then $\chi_0 = D_0$, so τ_n^* is regular and nonparametric efficient.

Consequently, $\tau_n^* - \tau_0 = P_n\chi_0 + o_p(n^{-1/2})$ and $n^{1/2}(\tau_n^* - \tau_0)$ converges in distribution to a mean-zero normal random variable with variance $\sigma_0^2 := \mathrm{var}_0\{\chi_0(Z)\}$.

Given a consistent estimator of σ_0^2, inference may be conducted using the bootstrap, as described in Section 2.5.3, or via Wald-type confidence intervals and tests based on the influence function χ_0. For Wald-type inference, the unknown quantities r_0 and s_0 in χ_0 can be estimated via empirical risk minimization, as discussed in Appendix 2.D.5. If both nuisance estimators are consistent, then r_0 and s_0 are identically zero, and standard confidence intervals based on the calibrated nuisances are valid without additional correction. For parameters that can be expressed as smooth transformations of linear summaries—such as relative differences between counterfactual means—DRAL estimators can be obtained by first constructing DRAL estimators of each summary and then applying the corresponding plug-in map. Inference may then be based on either the delta method or the bootstrap. For instance, in Example 2.1, DRAL estimators of the ATE $\tau_0(1) - \tau_0(0)$ and relative ATE $\tau_0(1)/\tau_0(0)$ are $\tau_n^*(1) - \tau_n^*(0)$ and $\tau_n^*(1)/\tau_n^*(0)$, respectively.

When both nuisance estimators are consistent, Theorem 2.4 shows that the isocalibrated DML estimator is regular, asymptotically linear, and efficient for τ_0 under weaker conditions than standard doubly robust estimators and related DRAL estimators. In particular, it avoids product-rate conditions and allows one nuisance function to be estimated arbitrarily slowly. By contrast, the TMLE-based DRAL estimators of the ATE proposed in van der Laan [2014] and related work [Benkeser et al., 2017, Dukes et al., 2024] typically require the nuisance estimators to converge sufficiently quickly to their (possibly incorrect) limits, imposing $\|\mu_{n,\diamond}^* - \overline{\mu}_0\|_{\overline{P}_0} \|\alpha_{n,\diamond}^* - \overline{\alpha}_0\|_{\overline{P}_0} = o_P(n^{-1/2})$. When both nuisance estimators are consistent, this offers no improvement over rate double robustness. We view this requirement as an artifact of existing theoretical analyses rather than an inherent limitation of TMLE-based

estimators, and our analysis can be adapted to relax it. For the expected conditional covariance parameter, Bonvini et al. [2024] recently showed that the estimation rate achieved by the isocalibrated DML estimator is minimax optimal under a hybrid model that combines smoothness and structure-agnostic assumptions on the nuisance functions, similar to the conditions of Lemma 2.4. Consequently, the rate conditions for asymptotic normality imposed by C1 cannot be meaningfully improved without additional assumptions.

The convergence-rate conditions in C1 are stated in terms of the isotonic-calibrated nuisance estimators, which raises the question of whether the calibration step can degrade the rates achieved by the first-stage learners. Appendix 2.D.4 addresses this by showing that isotonic calibration preserves the L^2 convergence rates of the first-stage nuisance estimators, up to an additional $O_p(n^{-1/3})$ term. Since $n^{-1/3} = o(n^{-1/4})$, this calibration error is negligible relative to the $o_p(n^{-1/4})$ rate required in C1, and thus calibration imposes no additional rate restrictions. Moreover, in the high-dimensional and nonparametric regimes that motivate our work, the one-dimensional $n^{-1/3}$ calibration term is typically dominated by first-stage estimation error. This term reflects estimation of a monotone function in one dimension and therefore depends favorably on problem constants, independent of covariate dimension [Chatterjee et al., 2013]. By contrast, rates faster than $n^{-1/3}$ for high-dimensional first-stage learning typically require strong smoothness or extreme sparsity assumptions, and can involve constants that scale poorly with dimension and other structural parameters.Lemma 2.12 in Appendix 2.D.4 further shows that the calibrated nuisance estimators achieve the same mean-squared-error rate as the best monotone transformation of the original nuisance estimators, where "best" is defined as the minimizer of the relevant population risk: least-squares risk for the outcome regression and Riesz risk for the Riesz representer. Thus, isotonic calibration relaxes the consistency requirement by allowing the initial nuisance estimators to be misspecified up to an unknown monotone transformation.

When $\tau_0 = E_0\{\mu_0(1, W)\}$ is a counterfactual mean, the DRAL properties in Lemma 2.5 have a simple interpretation: confounding bias can be removed by adjusting either for the outcome regression $\mu_0(A, W)$ or for the inverse-propensity weight $\alpha_0(A, W) = 1\{A = 1\}/P_0(A = 1 \mid W)$. The latter amounts to adjusting for the propensity score $P_0(A = 1 \mid W)$, a classical balancing score [Rosenbaum and Rubin, 1983, Imai and Ratkovic, 2014], whereas

$\mu_0(A, W)$ is a deconfounding score [Benkeser et al., 2020, D'Amour and Franks, 2021]. More formally, recalling the notation of Section 2.4.1, the calibration property in (2.12) debiases the plug-in estimator $\psi_n(\mu_n)$ under models in which $\mu_n(A, W)$ provides a dimension reduction of (A, W) [Benkeser et al., 2020]. Similarly, the calibration property in (2.13) reduces to covariate balance conditional on the estimated propensity score [van der Laan et al., 2024b], and thus debiases the weighted-outcome estimator $\frac{1}{n}\sum_{i=1}^{n}\alpha_n(A_i, W_i)Y_i$ under models in which $\alpha_n(A, W)$ provides a dimension reduction [Rosenbaum and Rubin, 1983]. Intuitively, the calibrated DML estimator combines these two debiasing mechanisms, yielding bias correction whenever at least one nuisance estimator is consistent.

The calibrated IPW and plug-in estimators are special cases of our general framework, obtained by setting $\mu_{n,j}$ or $\alpha_{n,j}$ to zero, respectively. By Theorem 2.4, each is asymptotically linear. However, these singly robust estimators will not generally be regular and efficient even under correct specification, and so we do not recommend using them in practice.

Irregularity under nuisance misspecification

Under the stated conditions, Theorem 2.4 shows that calibrated DML is regular whenever both nuisance estimators are consistent and at least one converges faster than $n^{-1/4}$. When flexible learners are used, this consistency requirement is mild. Indeed, many machine learning methods—including neural networks, random forests, and gradient boosting—are universal approximators and can consistently estimate smooth nuisance functions under suitable conditions [Hornik et al., 1989, Anthony and Bartlett, 2009, Zhang and Yu, 2005, Biau et al., 2008]. Calibrated DML and related DRAL estimators [Benkeser et al., 2017, Dukes et al., 2024, Bonvini et al., 2024] may be *irregular* under nuisance misspecification, that is, when one nuisance estimator is inconsistent. Regularity is desirable because it enables uniformly valid asymptotic inference; without it, there may exist distributions for which the asymptotic approximation requires arbitrarily large sample sizes to be accurate [Leeb and Pötscher, 2005]. Nevertheless, when one nuisance function is inconsistent, the estimator remains asymptotically normal and continues to support pointwise valid inference, with the resulting irregularity quantified in Appendix 2.E. There, we show that under any

fixed local perturbation $P_{0,t/\sqrt{n}}$ of P_0, the rescaled estimation error $\sqrt{n}\{\tau_n^* - \Psi(P_{0,t/\sqrt{n}})\}$ is asymptotically normal with a nonzero mean of order $t(\|\overline{\mu}_0 - \mu_0\| \vee \|\overline{\alpha}_0 - \alpha_0\|)$. Consequently, confidence intervals for calibrated DML estimators may still attain close-to-nominal coverage provided that either the magnitude of perturbation ($|t|$) or misspecification $\big(\|\overline{\mu}_0 - \mu_0\| \vee \|\overline{\alpha}_0 - \alpha_0\|\big)$ is sufficiently small.

2.5.3 Doubly robust validity of bootstrap-assisted confidence intervals

We now show the bootstrap-assisted confidence intervals from Algorithm 2 provide doubly robust asymptotically valid inference for τ_0. Like the Wald interval based on Theorem 2.4, constructing our bootstrap-assisted intervals does not require having prior knowledge of which nuisance estimator is consistent, or whether both are. However, unlike the Wald approach, it does not require estimating the unknown quantities r_0 and s_0 appearing in the influence function of τ_n^*. Hence, this bootstrap procedure is fully automated, only requiring computing the isocalibrated DML estimator within each bootstrap sample.

We recall that Algorithm 2 assumes ψ_0 has the form $\mu \mapsto E_0\{m(Z, \mu)\}$ and $\frac{1}{J}\sum_{j=1}^J \psi_{n,j}$ is taken to be the empirical plug-in estimator $\mu \mapsto \frac{1}{n}\sum_{i=1}^n m(Z_i, \mu)$. Below, we define $P_{n,j}^{\#}$ as the empirical distribution of an arbitrary bootstrap sample $\mathcal{C}^{\#(j)}$ drawn from the empirical distribution $P_{n,j}$ of $\mathcal{C}^{(j)}$. Moreover, we denote the bootstrapped calibrated nuisance estimators corresponding to $P_{n,j}^{\#}$ as $\mu_{n,j}^{\#}$ and $\alpha_{n,j}^{\#}$, which are obtained as in Algorithm 2. Finally, we define $\tau_n^{\#}$ as the corresponding bootstrapped isocalibrated DML estimator obtained from $\mu_{n,1}^{\#}, \ldots, \mu_{n,J}^{\#}$ and $\alpha_{n,1}^{\#}, \ldots, \alpha_{n,J}^{\#}$. Denoting by $\Delta_{\alpha,n,j}^{\#} := \alpha_{n,j}^{\#} - \alpha_0$ and $\Delta_{\mu,n,j}^{\#} := \mu_{n,j}^{\#} - \mu_0$ for $j = 1, \ldots, J$ the nuisance estimation errors, the following conditions are used in the upcoming theorem:

D1) *Boundedness of bootstrap nuisance estimators:* The functions $\mu_{n,j}^{\#}$ and $\alpha_{n,j}^{\#}$ are uniformly bounded by M with P_0–probability tending to one for each $j = 1, \ldots, J$;

D2) *Partial orthogonality and projection error bounds:* For each $j \in [J]$, assume that the following holds jointly with whichever part of C1 is assumed:

i) Under C1i, $\overline{\gamma}_{\alpha,0}(\mu^{\#}_{n,\diamond}, \alpha^{\#}_{n,\diamond}) = O_p(\|\alpha^{\#}_{n,\diamond} - \alpha_0\|^2_{\overline{P}_0})$ and $\|(\overline{\Pi}_{\alpha^{\#}_{n,\diamond}} - \overline{\Pi}_{\alpha_0})\Delta^{\#}_{\mu,n,\diamond}\|_{\overline{P}_0} = o_p(1)$.

ii) Under C1ii, $\overline{\gamma}_{\mu,0}(\mu^{\#}_{n,\diamond}, \alpha^{\#}_{n,\diamond}) = O_p(\|\mu^{\#}_{n,\diamond} - \mu_0\|^2_{\overline{P}_0})$ and $\|(\overline{\Pi}_{\mu^{\#}_{n,\diamond}} - \overline{\Pi}_{\mu_0})\Delta^{\#}_{\alpha,n,\diamond}\|_{\overline{P}_0} = o_p(1)$.

D3) *Pointwise Lipschitz continuity of the functional:* $\mathcal{A}$ is finite, and there exists a constant $L \in (0, \infty)$ such that, for all $\mu \in \mathcal{H}$, $\|m(\cdot, \mu)\| \leq L\|\mu\|$ and $|m(Z, \mu)| \leq L \max_{a\in\mathcal{A}} |\mu(a, W)|$ P_0-almost surely.

Condition D3 is used to formally verify C4 and an analogous condition for the bootstrapped estimators.

Theorem 2.5 (Double robustness of confidence intervals)**.** *Under the conditions of Theorem 2.4 and D1–D3, the conditional law of* $n^{\frac{1}{2}}(\tau^{\#}_n - \tau^*_n)$ *given* $\{P_{n,j} : j = 1, \ldots, J\}$ *converges marginally in* P_0*–probability to the law of a mean-zero normal random variable with variance* $\sigma^2_0 := var_0\{\chi_0(Z)\}$.

Under regularity conditions, as the number of bootstrap samples $K \to \infty$, the scaled version $n^{\frac{1}{2}}\sigma^{\#}_n$ of the bootstrap standard error of Algorithm 2 is a consistent estimator of the limiting standard deviation σ_0 of $n^{\frac{1}{2}}(\tau^*_n - \tau_0)$. Hence, by Theorem 2.5, the bootstrap-assisted confidence interval $\mathrm{CI}_n(\rho)$ from Algorithm 2 has doubly robust asymptotically exact coverage, in the sense that $P_0\{\tau_0 \in \mathrm{CI}_n(\rho)\} \to 1-\rho$ as $n \to \infty$, provided $K \to \infty$ sufficiently quickly.

2.6 Numerical experiments

2.6.1 Experiment 1: performance under inconsistent nuisance estimation

We evaluate the numerical performance of the calibrated DML estimator through simulation, focusing on inference for the ATE $\tau_0 = E_0\{\mu_0(1, W) - \mu_0(0, W)\}$. The estimator is defined as $\tau^*_n := \tau^*_n(1) - \tau^*_n(0)$, where $\tau^*_n(a)$ is the calibrated DML estimator of the counterfactual mean $E_0\{\mu_0(a, W)\}$, constructed as in Example 2.1 using Algorithm 1. We compare its

performance to the standard cross-fitted AIPW estimator from Section 2.2.1 and the DR-TMLE estimator from Benkeser et al. [2017], described in Appendix 2.D.6 and implemented in the `drtmle R` package [Benkeser and Hejazi, 2023]. To quantify uncertainty, we use bootstrap confidence intervals (Algorithm 2) for calibrated DML and Wald-type intervals based on influence function variance estimates for AIPW and DR-TMLE.

We consider the same data-generating process used in the experiments of Benkeser et al. [2017]. Covariates $W_1 \sim \text{Uniform}(-2, 2)$ and $W_2 \sim \text{Bernoulli}(0.5)$ are drawn independently. The treatment and outcome are binary, with propensity $\pi_0(1 \mid (w_1, w_2)) := \text{expit}(-w_1 + 2w_1w_2)$ and outcome regression $\mu_0(a, w_1, w_2) := \text{expit}(0.2a - w_1 + 2w_1w_2)$. To evaluate the doubly robust inference property, we examine three scenarios of inconsistent nuisance estimation for each estimator:

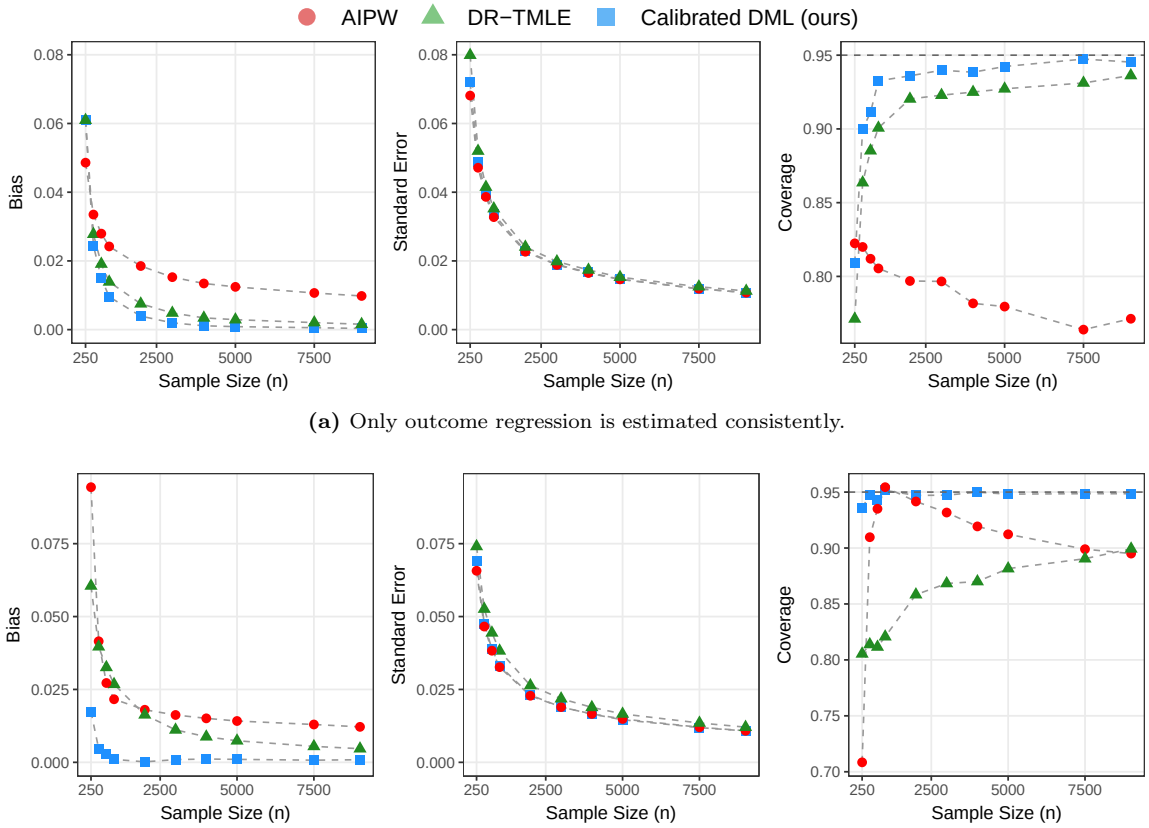
(a) when both the outcome regression and treatment mechanism are estimated consistently;

(b) when only the treatment mechanism is estimated consistently;

(c) when only the outcome regression is estimated consistently.

In settings (b) and (c), to induce inconsistency in the estimation of the outcome regression and treatment mechanisms, we estimate the propensity score π_0 and outcome regressions μ_0 using misspecified main-term logistic regression models. In all settings, any consistently estimated nuisance function is obtained using a 5-fold cross-fitted Nadaraya–Watson kernel smoother with bandwidth selected via cross-validation and stratified by both A and W_2 [Hofmeyr, 2020]. For DR-TMLE, we apply the same kernel method to estimate the nuisance functions r_0 and s_0. Further details and code for reproducing these analyses are available in the GitHub repository linked in Appendix 2.H.

For each scenario and sample sizes ranging from 250 to 10,000, we generated and analyzed 5,000 datasets. Empirical estimates of bias, standard error, and 95% confidence interval coverage are presented in Figure 2.3. Results for the setting in which both nuisance functions are consistently estimated appear in Appendix 2.I.1. In scenarios with inconsistent nuisance estimation, both calibrated DML and DR-TMLE exhibited DRAL behavior, showing lower



(a) Only outcome regression is estimated consistently.

(b) Only propensity score is estimated consistently.

Figure 2.3: Empirical bias, standard error, and 95% confidence interval coverage for calibrated DML (IC-DML), DR-TMLE and AIPW estimators under singly consistent estimation of the outcome regression and treatment mechanism.

bias and better coverage than AIPW. Relative to DR-TMLE, calibrated DML exhibited less bias, achieved better coverage, and had smaller or comparable variance. In settings where both nuisance functions were consistently estimated, all estimators performed similarly, as expected.

2.6.2 Experiment 2: benchmarking on semi-synthetic datasets

In this experiment, we evaluate the performance of our method on a collection of semi-synthetic datasets commonly used to benchmark causal inference methods. These datasets feature moderately high-dimensional covariates, limited overlap, confounding, and treatment

Table 2.2: Sample size of total study (n_{total}), expected size of treatment arm ($n_{treated}$), and covariate dimension (d) for the semi-synthetic datasets of second experiment. For the ACIC-2017 synthetic dataset, we employed simulation setting 24, characterized by additive errors, a large signal, large noise, and strong confounding.

Dataset	n_{total}	$n_{treated}$	d	source
ACIC-2017	4302	2233	58	Hahn et al. [2019]
ACIC-2018	varied	varied	varied	Shimoni et al. [2018]
IHDP	672	123	25	Shalit et al. [2017]
Lalonde CPS	16177	167	9	Neal et al. [2020]
Lalonde PSID	2675	175	9	Neal et al. [2020]
Twins	11984	8935	76	Neal et al. [2020]

imbalance (Table 2.2). We include the Twins, Lalonde PSID, and Lalonde CPS datasets from the *realCause* generative modeling framework [Neal et al., 2020], which are derived from original data sources [Louizos et al., 2017, LaLonde, 1986]. We also use several datasets from the 2017 Atlantic Causal Inference Challenge (ACIC-2017, settings 17–24) [Hahn et al., 2019], which are based on covariates from the Infant Health and Development Program (IHDP) [Brooks-Gunn et al., 1992], and 680 datasets from the scaling component of ACIC-2018 [Shimoni et al., 2018], with sample sizes of 1,000, 2,500, 5,000, and 10,000. The final dataset (IHDP) is a semi-synthetic version of the IHDP data introduced in Hill [2011] (Setting B) and further analyzed in Shalit et al. [2017]. In these benchmarks, the covariates are fixed and derived from real data, and the inferential target is the sample average treatment effect, defined as $\tau_{n,0} := \frac{1}{n}\sum_{i=1}^{n}(Y_{1,i} - Y_{0,i})$, the empirical mean of the individual treatment effects in the observed sample.

We compare the calibrated DML estimator to the AIPW estimator, with both defined as in the previous section. DR-TMLE was excluded from this experiment due to the computational expense of the iterative debiasing step, as discussed in the Introduction. Cross-fitted nuisance functions (propensity score and outcome regression) are estimated using gradient-boosted regression trees via `xgboost` [Chen et al., 2015], with the maximum tree depth selected by 10-fold cross-validation. Under smoothness conditions, these nuisance estimators will be consistent, though the rate of convergence may be slow in moderate dimensions [Zhang and Yu, 2005]. For the AIPW estimator, we truncate the cross-fitted propensity score estimates to lie in $[c_n, 1 - c_n]$, where $c_n := \min\{0.05, 25/(n^{1/2}\log n)\}$, following Gru-

Table 2.3: Evaluation of Absolute Bias, Root Mean Square Error (RMSE), and 95% Confidence Interval coverage on semi-synthetic benchmark datasets. Both outcome regression and propensity scores are estimated using an ensemble of gradient-boosted trees. For comparability across rows, the absolute bias and RMSE is scaled by a constant equal to the average effect size across dataset realizations for that benchmark. For clarity, we report only sample size 10000 results for ACIC-2018 and the settings with strong confounding for ACIC-2017. The full results can be found in Appendix 2.I.2.

Dataset	**Coverage**		**RMSE**		**Absolute Bias**	
	AIPW	C-DML	AIPW	C-DML	AIPW	C-DML
ACIC-2017 (18)	0.27	**0.64**	0.70	**0.58**	**0.21**	0.28
ACIC-2017 (20)	0.32	**0.90**	2.00	**1.40**	1.60	**0.20**
ACIC-2017 (22)	0.56	**0.81**	0.11	**0.10**	**0.004**	0.035
ACIC-2017 (24)	0.32	**0.90**	0.34	**0.25**	0.30	**0.04**
ACIC-2018 (10000)	0.55	**0.68**	690	**680**	17	**13**
IHDP	**0.57**	**0.57**	**0.46**	**0.46**	**0.13**	**0.13**
Lalonde CPS	0.16	**0.75**	**0.22**	0.34	0.14	**0.084**
Lalonde PSID	0.46	**0.84**	**0.19**	0.44	**0.04**	**0.04**
Twins	0.51	**0.54**	0.24	**0.23**	0.22	**0.21**

ber et al. [2022]. We do not apply additional truncation for calibrated DML, since isotonic calibration induces automatic truncation [van der Laan et al., 2024b].

For each collection of k dataset realizations, Table 2.3 reports empirical absolute bias, root mean square error (RMSE), and nominal 95% confidence interval coverage. The calibrated DML estimator consistently achieved superior coverage performance. For instance, in the ACIC-2017 settings, it improved AIPW's coverage rates from 27% to 64%, 32% to 90%, 56% to 81%, and 32% to 90%. In terms of point estimation, calibrated DML also generally exhibited lower or comparable absolute bias across datasets. Notably, in the ACIC-2017 settings 20 and 24, the calibrated DML estimator outperformed the AIPW estimator, achieving bias reductions by factors of approximately 8. In ACIC-2017 settings 24 and 20, the calibrated DML estimator achieved lower RMSEs than AIPW (0.25 vs. 0.35 and 1.4 vs. 2.0, respectively). While AIPW performs better in terms of RMSE in the Lalonde CPS and Lalonde PSID datasets, this came at the cost of substantially poorer coverage.

Overall, the calibrated DML estimator outperformed the AIPW estimator in terms of confidence interval coverage, RMSE, and absolute bias across a range of semi-synthetic datasets.

2.7 Conclusion

In this work, we introduced a unified machine learning framework for doubly robust inference on linear functionals of the outcome regression. These functionals belong to a subset of parameters characterized by the mixed bias property [Rotnitzky et al., 2021]. For the case of linear functionals, our framework provides an affirmative answer to the open question raised by Smucler et al. [2019], regarding the feasibility of achieving learner-agnostic doubly robust inference for mixed-bias parameters using black-box nuisance function estimators. Another significant contribution of this work lies in establishing a close relationship between calibration and the doubly robust inference properties of debiased machine learning estimators. Importantly, our proof techniques rely solely on the nuisance estimators satisfying the calibration score equations stated in Section 2.5.2. While we applied our calibrated DML framework with isotonic calibration, these score equations can also be satisfied when nuisance functions are calibrated using other histogram-type estimators, such as honest regression trees [Wager and Athey, 2018], the highly adaptive lasso [van der Laan, 2017, Benkeser and Van Der Laan, 2016, Fang et al., 2021], and Venn-Abers calibration [Vovk and Petej, 2012, van der Laan and Alaa, 2024, 2025]. In general, both our theoretical and empirical results support the use of calibrated nuisance estimators in causal inference to mitigate bias arising from inconsistent nuisance estimation and to ensure calibrated inferences.

There remain several open questions along this line of research. First, while we outline an extension of our framework to general parameters characterized by the mixed bias property [Rotnitzky et al., 2021] in Appendix 2.F.1, we leave the full development of this approach to future work. Using this extension, we propose calibrated DML estimators for the expected conditional covariance parameter, which arises in partially linear regression models, analogous to those proposed by Dukes et al. [2024] and Bonvini et al. [2024]. Second, it may be of interest to develop an isotonic calibrated targeted minimum loss estimator that inherits both the plug-in property and doubly robust asymptotic linearity, as considered in van der Laan [2014] and Benkeser et al. [2017]. Such an estimator could outperform non-plug-in alternatives such as one-step debiased estimators [Porter et al., 2011]. The main challenge is that isotonic calibration and TMLE targeting enforce different (and not generally compatible)

score equations, so applying one step can undo the constraints imposed by the other. A potential approach for this could be to iteratively apply the isotonic calibration and targeting steps until the relevant score equations for both DRAL and the plug-in property are solved. Alternatively, the score-preserving TMLE of Pimentel et al. [2025] could be applied to the calibrated nuisance estimates to ensure that the targeting step preserves the empirical calibration scores. Extending our framework to accommodate multiple and sequentially robust estimators in longitudinal data structures presents further exciting research opportunities [Rotnitzky et al., 2017, Luedtke et al., 2017].

As discussed in the Introduction, the approach of Liu et al. [2023] to constructing DRAL estimators in high-dimensional generalized linear models can be viewed as a sparsity-based analogue of our calibration-based post-hoc correction approach. Unlike ours, their method relies on first-stage ℓ_1-regularized estimators, a high-dimensional second-stage adjustment, and sparsity-based conditions, whereas our work is learner-agnostic, uses a one-dimensional adjustment, and is not tied to ℓ_1-regularization or sparsity assumptions. An interesting open question is whether the first-stage estimators in Liu et al. [2023] can be replaced by user-supplied machine learners rather than being restricted to ℓ_1-regularized estimators in a high-dimensional linear model. In Appendix 2.F.2, we extend our calibration framework and Theorem 2.3 to more general orthogonality conditions and projection operators, as occur in regularized empirical risk minimization and sieve estimation. This extension suggests one possible route toward generalizing Liu et al. [2023] to flexible first-stage learners, with assumptions stated as sparsity conditions on first-stage nuisance *errors* rather than on the nuisance *functions* themselves. More broadly, it suggests that both calibration and multi-accuracy, viewed as post-hoc predictive-modeling techniques, may be used to construct DRAL estimators. These connections point toward a more unified view of sparsity-based doubly robust methods, including Liu et al., 2023, our calibration-based approach, and TMLE-style methods, including Benkeser et al., 2017.

Acknowledgements. Research reported in this publication was supported by NIH grants DP2-LM013340 and R01-HL137808, and NSF grant DMS-2210216. The content is solely the responsibility of the authors and does not necessarily represent the official views of the

funding agencies.

Chapter Appendix

2.A Additional discussion of sparsity-based doubly robust inference

This section expands the comparison in Section 2.1.3 between our calibration-based approach and sparsity-based doubly robust asymptotic linearity methods, especially Liu et al. [2023].

When nuisances are estimated via ℓ_1-regularized empirical risk minimization over linear classes, the Karush–Kuhn–Tucker (KKT) conditions imply approximate empirical orthogonality over suitable linear subspaces (see, e.g., Tan, 2020, Eq. 15; Bradic et al., 2019b, Sec. 2.1; Smucler et al., 2019, Sec. 4.1). Under additional approximation conditions—typically requiring at least one nuisance error to be sufficiently sparse—this yields doubly robust asymptotic linearity for ℓ_1-regularized autoDML estimators. Liu et al. [2023] show that this type of ℓ_1-regularized, high-dimensional correction can be applied post hoc to initial ℓ_1-regularized learners. For the outcome regression, their post-hoc correction can be viewed as enforcing a high-dimensional orthogonality objective—often called *multi-accuracy* [Kim et al., 2019, Deng et al., 2023, Roth, 2022]—by targeting the moment conditions $E_0\left[(Y-\mu_n(A,W))\, f(A,W)\right]=0$ for all f in a rich class (in their case, the span of the ℓ_1 design/basis functions).

In contrast, Section 2.3.2 shows that, for our purpose, it suffices to target a single moment condition, namely $E_0\left[(Y-\mu_n(A,W))\, s_{n,0}(A,W)\right]=0$, as in (2.4). Our approach, calibrated DML (Section 2.4.1), instead targets the one-dimensional calibration constraint $\mu_n(A,W)=E_0\left[Y \mid \mu_n(A,W)\right]$, or, equivalently, the one-dimensional orthogonality conditions $E_0\left[(Y-\mu_n(A,W))\,\theta\big(\mu_n(A,W)\big)\right]=0$ for all $\theta:\mathbb{R}\to\mathbb{R}$.

2.B Additional discussion of conditions

2.B.1 Discussion of Condition A2

Hellinger continuity of the map $P\mapsto\psi_P$ at P_0 in Condition A2 is a mild regularity requirement. The following lemma shows that this condition holds for expectation-representable linear functionals in Example 2.2.

Lemma 2.6 (Hellinger continuity of expectations)**.** *Let P and P' be probability measures on $(\mathcal{Z},\mathcal{A})$ that are dominated by a common σ-finite measure ν, with densities p and p'. For*

any measurable $f : \mathcal{Z} \to \mathbb{R}$ *such that* $E_P[f(Z)^2] < \infty$ *and* $E_{P'}[f(Z)^2] < \infty$*, we have*

$$\big|E_P[f(Z)] - E_{P'}[f(Z)]\big| \;\le\; 2\,H(P,P')\left(E_P[f(Z)^2] + E_{P'}[f(Z)^2]\right)^{1/2},$$

where $H(P,P') := \left(\frac{1}{2}\int(\sqrt{p}-\sqrt{p'})^2\,d\nu\right)^{1/2}$ *denotes the Hellinger distance.*

Proof deferred to Appendix A.

2.C Discussion of conditions of Theorem 2.4

Condition C1 is the key assumption underlying the DRAL behavior of the calibrated DML estimator. Condition C2 ensures that the expansion in Theorem 2.3 applies to the calibrated nuisance estimators; sufficient conditions are given in Lemma 2.4. The associated projection error bounds also control the empirical process remainders. These extra conditions are not needed under the standard cross-product rate $\|\mu^*_{n,j} - \mu_0\|\,\|\alpha^*_{n,j} - \alpha_0\| = o_p(n^{-1/2})$, in which case asymptotic linearity and efficiency of isocalibrated DML follow from conventional arguments.

Condition C4 appears because the linear functional must also be estimated. It holds whenever both (i) $\psi_{n,j}(f_{n,j}) - \psi_0(f_{n,j}) - P_{n,j}\phi_{0,f_{n,j}} = o_p(n^{-1/2})$ and (ii) $(P_{n,j} - P_0)\{\phi_{0,f_{n,j}} - \phi_{0,f_0}\} = o_p(n^{-1/2})$. Condition (i) requires the linearization remainder of $\psi_{n,j}(r^*_{n,j})$ as an estimator of $\psi_0(r^*_{n,j})$ to be asymptotically negligible. This condition is automatic for functionals of the form in (2.2), that is, for $\psi_0 : \mu \mapsto E_0\{m(Z,\mu)\}$ for some map $(z,\mu) \mapsto m(z,\mu)$, when $\frac{1}{J}\sum_{j=1}^J \psi_{n,j}$ is taken to be the empirical plug-in estimator $\mu \mapsto \frac{1}{n}\sum_{i=1}^n m(Z_i,\mu)$. Condition (ii) is a mild empirical process requirement that can be verified with the same proof techniques used below. For example, when $\mathcal{A}$ is finite, (ii) holds under the Lipschitz continuity condition

$$|\phi_{0,\mu^*_{n,j}}(z) - \phi_{0,\overline{\mu}_0}(z)| \le L \sup_{a_0\in\mathcal{A}} |\mu^*_{n,j}(a_0,w) - \overline{\mu}_0(a_0,w)|, \quad P_0\text{-a.e.}$$

for some $L < \infty$, together with the convergence condition

$$\sup_{a_0\in\mathcal{A}} \|\mu^*_{n,j}(a_0,\cdot) - \overline{\mu}_0(a_0,\cdot)\| = o_p(1),$$

see Lemmas A.8 and A.13 in the Appendix.

2.D Additional details on DRAL and calibrated DML

This section collects examples and auxiliary discussion that support the general framework of Sections 2.3.2 and 2.4.1. The subsections below provide additional examples, conditions used in the main theory, and implementation details for the calibrated DML procedure.

2.D.1 Additional examples of parameters

Example 2.7 (counterfactual means under general interventions)**.** For simplicity, suppose $\mathcal{A}$ is discrete, and denote the propensity score by $\pi_0(a \mid w) := P_0(A = a \mid W = w)$. Let Y^{a_0} denote the potential outcome that would be observed if treatment $a_0 \in \mathcal{A}$ were administered. Under causal conditions [Rubin, 1974], the mean of Y^{a_0} is identified by $\tau_0 := E_0\{\mu_0(a_0, W)\}$, which corresponds to the linear functional $\psi_0 : \mu \mapsto E_0\{\mu(a_0, W)\}$. Doubly robust inference for this parameter was considered in van der Laan [2014] and Benkeser et al. [2017]. More generally, the mean of the counterfactual outcome under a stochastic intervention that, given $W = w$, draws a treatment value A^* at random from the distribution with probability mass function $a \mapsto \omega_0(a|w)$ depending only on $P_{0,A,W}$ can be identified by $\tau_0 := \int\int \mu_0(a, w)\omega_0(a|w)P_{0,A,W}(da, dw)$. If $(a, w) \mapsto \alpha_0(a, w) := \omega_0(a|w)/\pi_0(a|w)$ has finite $\mathcal{H}$–norm, then conditions A1–A3 hold and α_0 is the nonparametric Riesz representer of τ_0.

Example 2.8 (Expected conditional covariance)**.** Consider the expected conditional covariance between A and Y adjusted for W,

$$\tau_0 = E_0[\{A - \pi_0(W)\}\{Y - m_0(W)\}],$$

where $\pi_0(w) := E_0[A \mid W = w]$ and $m_0(w) := E_0[Y \mid W = w]$. This parameter arises in semiparametric ATE estimation under partially linear models [Robinson, 1988, Li et al., 2011], as a measure of "causal influence" [Díaz, 2024], and in conditional independence testing [Shah and Peters, 2020]. Note that $\tau_0 = E_0[AY] - E_0\{A\, m_0(W)\}$. The first term is efficiently estimated by the empirical plug-in $n^{-1}\sum_{i=1}^n A_i Y_i$. The second is a linear functional of m_0, namely $\psi_0 : m \mapsto E_0\{A\, m(W)\}$, with Riesz representer π_0, and can be estimated by its empirical plug-in $\psi_n : m \mapsto n^{-1}\sum_{i=1}^n A_i m(W_i)$. Consequently, $E_0\{A\, m_0(W)\}$ is an estimand within our framework, and our methods apply to doubly robust inference for τ_0.

Technically, ψ_0 depends on the marginalized conditional mean m_0 rather than the treatment-specific conditional mean μ_0; however, this does not affect the applicability of our approach, since m_0 can be calibrated under squared error loss in the same way (see Appendix 2.F.1). Related work on doubly robust inference for this parameter includes Dukes et al. [2024] and Bonvini et al. [2024].

Example 2.9 (average treatment effect under covariate-dependent outcome missingness)**.** Suppose that the observed data unit consists of $(W, S, \Delta, \Delta Y)$, where $S \in \{0, 1\}$ is a binary treatment, $\Delta \in \{0, 1\}$ is a missingness indicator taking value 1 if Y is observed, and ΔY is the observed outcome, equal to Y if $\Delta = 1$ and 0 otherwise. Under causal conditions, the ATE under covariate-dependent outcome missingness is identified by the estimand $\tau_0 := E_0\{E_0(\Delta Y \,|\, S = 1, \Delta = 1, W) - E_0(\Delta Y \,|\, S = 0, \Delta = 1, W)\}$. Doubly robust inference for this estimand was considered in Díaz and van der Laan [2017]. Using the notation of this section, we can write $Z := (W, A, Y')$, where $A := (S, \Delta)$ is a multi-valued interventional variable and $Y' := \Delta Y$ is an outcome with missingness. The linear functional corresponding to τ_0 is $\psi_0 : \mu \mapsto E_0\{\mu((1, 1), W) - \mu((0, 1), W)\}$. The nonparametric Riesz representer of ψ_0 is given by $\alpha_0 : (s, \delta, w) \mapsto (2s - 1)\delta / P_0(\Delta = 1, S = s \,|\, W = w)$.

2.D.2 Controlling empirical process remainders in Theorem 2.3

In this section, we provide sufficient conditions for the asymptotic negligibility of the empirical-process remainders $(P_n - P_0)(A_{n,0} - A_0)$ and $(P_n - P_0)(B_{n,0} - B_0)$ that appear in the regression-based and Riesz-based decompositions in Theorem 2.3, respectively. Importantly, the lemmas show that these remainders are $o_p(n^{-1/2})$ under conditions requiring only that the nuisance associated with the decomposition is consistent, while the potentially inconsistent nuisance merely converges in probability to a limit. The proofs of the lemmas are given in Appendix A.2.4.

The next lemma establishes negligibility of $(P_n - P_0)(A_{n,0} - A_0)$. Recall that $A_0(z) = \Pi_{\mu_0}(\alpha_0 - \overline{\alpha}_0)\{y - \mu_0(z)\}$ and $A_{n,0}(z) = \Pi_{\mu_n}(\alpha_0 - \alpha_n)\{y - \mu_n(z)\}$.

(G1) *Boundedness.* μ_n, α_n, μ_0, α_0, and Y are uniformly bounded.

(G2) *Consistency and projection coupling.* $\|\overline{\alpha}_0 - \alpha_n\| = o_p(1)$, $\|\mu_n - \mu_0\| = o_p(1)$, and $\|\Pi_{\mu_n}(\overline{\alpha}_0 - \alpha_0) - \Pi_{\mu_0}(\overline{\alpha}_0 - \alpha_0)\| = O_p(\|\mu_n - \mu_0\|^\beta)$ for some $\beta \in (0, 1]$.

(G3) *Donsker condition.* With probability tending to one, $A_{n,0} - A_0$ belongs to a fixed Donsker function class.

Lemma 2.10. *Assume G1-G2. Then,* $\|A_{n,0} - A_0\| = o_p(1)$. *If, in addition, G3 holds, then* $(P_n - P_0)(A_{n,0} - A_0) = o_p(n^{-1/2})$.

In our proof of Theorem 2.4, to establish DRAL for the isocalibrated DML estimator with cross-fitting, we avoid Donsker assumptions such as G3. Instead, we use cross-fitting to remove any complexity conditions on the initial nuisance estimators $\alpha_{n,j}$ and $\mu_{n,j}$. We then exploit that the calibrated nuisance estimators $\alpha^*_{n,j}$ and $\mu^*_{n,j}$ differ from $\alpha_{n,j}$ and $\mu_{n,j}$ only through monotone transformations. Since the class of monotone functions has finite uniform entropy integral (and is therefore Donsker), this controls the complexity of the calibrated nuisance estimators. With cross-fitting, to control the complexity of functions of the form $s_{n,0} = \Pi_{\mu_n}\Delta_{\alpha,n}$, it suffices that the transformation $\theta_{\alpha,n} : \mathbb{R} \to \mathbb{R}$ defined by $s_{n,0} = \theta_{\alpha,n} \circ \mu_n$ has bounded variation (as in C6), which likewise yields controlled complexity. See our proofs in Appendix A.3.5 for details.

The next lemma establishes negligibility of $(P_n - P_0)(B_{n,0} - B_0)$. Recall that $B_0 := \phi_{r_0} + \psi_0(r_0) - r_0\alpha_0$, and $B_{n,0} := \phi_{r_{n,0}} + \psi_0(r_{n,0}) - r_{n,0}\alpha_n$. Note that

$$(P_n - P_0)(B_{n,0} - B_0) = (P_n - P_0)\{\phi_{r_{n,0}} - \phi_{r_0} - r_{n,0}\alpha_n + r_0\alpha_0\}.$$

(G4) *Consistency and projection coupling.* $\|\overline{\mu}_0 - \mu_n\| = o_p(1)$, $\|\alpha_n - \alpha_0\| = o_p(1)$, and $\|\Pi_{\alpha_n}(\overline{\mu}_0 - \mu_0) - \Pi_{\alpha_0}(\overline{\mu}_0 - \mu_0)\| = O_p(\|\alpha_n - \alpha_0\|^\beta)$ for some $\beta \in (0, 1]$.

(G5) *Lipschitz continuity*: $\|\phi_{r_{n,0}} - \phi_{r_0}\| = O_p(\|r_{n,0} - r_0\|)$.

(G6) *Donsker condition.* With probability tending to one, $\phi_{r_{n,0}} - \phi_{r_0} - r_{n,0}\alpha_n + r_0\alpha_0$ belongs to a fixed Donsker function class.

Lemma 2.11. *Assume G1–G5. Then,* $\|\phi_{r_{n,0}} - \phi_{r_0} - r_{n,0}\alpha_n + r_0\alpha_0\| = o_p(1)$. *If, in addition, G6 holds, then* $(P_n - P_0)(B_{n,0} - B_0) = o_p(n^{-1/2})$.

2.D.3 Bootstrap-assisted confidence intervals with cross-fitting

Algorithm 2 Bootstrap-assisted doubly robust CIs

Input: cross-fitted estimators $(\alpha_{n,j}, \mu_{n,j})$ excluding fold $\mathcal{C}^{(j)}$ for $j = 1, \ldots, J$, isocalibrated DML estimator τ_n^*; confidence level $1 - \rho$; bootstrap replication number K;

1: **for** $k = 1, \ldots, K$ **do**

2: for $s = 1, \ldots, J$, draw bootstrap sample $\mathcal{C}_k^{(s)\#}$ from $\mathcal{C}_k^{(s)}$;

3: set $\mathcal{D}_{n,k}^{\#} := \{Z_1^{\#}, \ldots, Z_n^{\#}\} = \{\mathcal{C}_k^{(s)\#} : s = 1, \ldots, J\}$;

4: for $s = 1, \ldots, J$, set $j(i) = s$ for each $i \in \{1, \ldots, n\}$ such that $Z_i^{\#} \in \mathcal{C}_k^{(s)\#}$;

5: compute calibrators $f_n^{\#(k)}$ and $g_n^{\#(k)}$ using pooled out-of-fold estimates as

$$f_n^{\#(k)} \in \operatorname*{argmin}_{f \in \mathcal{F}_{\text{iso}}} \sum_{i=1}^{n} \left\{ Y_i^{\#} - f(\mu_{n,j(i)}(A_i^{\#}, W_i^{\#})) \right\}^2;$$

$$g_n^{\#(k)} \in \operatorname*{argmin}_{g \in \mathcal{F}_{\text{iso}}} \sum_{i=1}^{n} \left\{ g(\alpha_{n,j(i)}(A_i^{\#}, W_i^{\#}))^2 - 2m(Z_i^{\#}, g \circ \alpha_{n,j(i)}) \right\};$$

6: for $s = 1, \ldots, J$, set $\mu_{n,s}^{(k)\#} := f_n^{(k)\#} \circ \mu_{n,s}$ and $\alpha_{n,s}^{(k)\#} := g_n^{(k)\#} \circ \alpha_{n,s}$;

7: compute calibrated one-step debiased estimator

$$\tau_{n,k}^{\#} := \frac{1}{n} \sum_{i=1}^{n} \left[m(Z_i^{\#}, \mu_{n,j(i)}^{(k)\#}) + \alpha_{n,j(i)}^{(k)\#}(A_i^{\#}, W_i^{\#}) \left\{ Y_i^{\#} - \mu_{n,j(i)}^{(k)\#}(A_i^{\#}, W_i^{\#}) \right\} \right];$$

8: **end for**

9: set $\sigma_n^{\#2} := \frac{1}{K} \sum_{k=1}^{K} (\tau_{n,k}^{\#} - \overline{\tau}_n^{\#})^2$ with $\overline{\tau}_n^{\#} := \frac{1}{K} \sum_{k=1}^{K} \tau_{n,k}^{\#}$;

10: construct interval $\mathrm{CI}_n(\rho) := (\tau_n^* - q(\rho)\sigma_n^{\#}, \tau_n^* + q(\rho)\sigma_n^{\#})$;

11: **return** *bootstrap-assisted confidence interval* $\mathrm{CI}_n(\rho)$ for τ_0.

2.D.4 Convergence rates of isotonic calibrated nuisances

The convergence rate conditions in Theorem 2.4 are stated for the calibrated nuisance estimators rather than the original ones. This raises a natural question: can the calibration step worsen the convergence rate? The following informal lemma summarizes the formal rate result proved in Section A.4. Up to a typically negligible error of order $O_p(n^{-2/3})$, the isotonic-calibrated estimators perform nearly as well in mean squared error as the best monotone non-decreasing transformation of the uncalibrated, cross-fitted nuisance estimators.

Lemma 2.12. *Suppose that ψ_0 has the form (2.2) with mapping $m : \mathcal{A} \times \mathcal{W} \times \mathcal{H} \to \mathbb{R}$, and that $\frac{1}{J} \sum_{j=1}^{J} \psi_{n,j}$ is taken to be the plug-in estimator $\mu \mapsto \frac{1}{n} \sum_{i=1}^{n} m(Z_i, \mu)$. Under the*

conditions of Theorem A.15, it holds that

$$\sum_{j=1}^{J} \|\mu^*_{n,j} - \mu_0\|^2 = \min_{f \in \mathcal{F}_{\text{iso}}} \sum_{j=1}^{J} \|f \circ \mu_{n,j} - \mu_0\|^2 + O_p(n^{-\frac{2}{3}}) ;$$
$$\sum_{j=1}^{J} \|\alpha^*_{n,j} - \alpha_0\|^2 = \min_{g \in \mathcal{F}_{\text{iso}}} \sum_{j=1}^{J} \|g \circ \alpha_{n,j} - \alpha_0\|^2 + O_p(n^{-\frac{2}{3}}) .$$

Since the identity transformation is itself monotone non-decreasing, the above lemma implies that the isotonic-calibrated nuisance estimators typically perform no worse than the uncalibrated ones, and may perform better when the uncalibrated estimators are poorly calibrated or only consistent up to a monotone transformation. Thus, beyond enabling doubly robust asymptotic linearity, isotonic calibration relaxes the consistency requirements on the nuisance estimators: it suffices that the initial cross-fitted nuisance estimators be consistent up to a monotone transformation.

2.D.5 Estimation of nuisance functions in limiting distribution

When calibrated DML is used together with the bootstrap-assisted inference procedure of Algorithm 2, it is not necessary to estimate the univariate nuisance functions $f_{0,\overline{\mu}_0}$ and $g_{0,\overline{\alpha}_0}$ that appear in the limiting distribution of Theorem 2.4. Nevertheless, these functions can be estimated by univariate empirical risk minimization if desired. For example, one may wish to estimate them when implementing the TMLE approach to doubly robust inference discussed below.

To estimate these functions, first observe that $f_{0,\overline{\mu}_0}$ is the Riesz representer of the linear functional $\alpha \mapsto \langle \alpha_0 - \overline{\alpha}_0, \alpha \rangle_{P_0} = \psi_0(\alpha) - \langle \overline{\alpha}_0, \alpha \rangle_{P_0}$ with respect to the linear space

$$\{f \circ \overline{\mu}_0 : f : \mathbb{R} \to \mathbb{R}, \|f \circ \mu\| < \infty\} .$$

Hence, by the Riesz representation stated above Theorem 2.1, $f_{0,\overline{\mu}_0}$ can be obtained via univariate empirical risk minimization after replacing $\overline{\mu}_0$ and $\overline{\alpha}_0$ with their estimators. Specifically, $f_{0,\overline{\mu}_0}$ can be estimated via offset univariate empirical risk minimization over $\{f \circ \mu^*_{n,\diamond} \in L^2(\overline{P}_{0,A,W})\}$ based on the empirical Riesz-loss criterion

$$f \mapsto \frac{1}{n} \sum_{i=1}^{n} \left[\left\{ \alpha^*_{n,j(i)} - (f \circ \mu^*_{n,j(i)}) \right\}^2 (A_i, W_i) - 2 \left\{ (f \circ \mu^*_{n,j(i)} \cdot \alpha^*_{n,j(i)})(A_i, W_i) - \psi_{n,j(i)} \left(f \circ \mu^*_{n,j(i)} \right) \right\} \right] .$$

Similarly, since $\overline{\mu}_{0,\alpha}(a,w) = E_0\left[Y - \overline{\mu}_0(A,W) | \alpha(A,W) = \alpha(a,w)\right]$, the function $g_{0,\overline{\alpha}_0}$ can be estimated by regressing the residuals $\{Y_i - \mu^*_{n,j(i)}(A_i,W_i)\}_{i\in[n]}$ on $\{\alpha^*_{n,j(i)}(A_i,W_i)\}_{i\in[n]}$ using, for example, locally adaptive regression splines [Mammen and van de Geer, 1997] or random forests [Breiman, 2001].

2.D.6 Doubly robust inference via TMLE

For the special case of doubly robust inference for the average treatment effect, van der Laan [2014] and Benkeser et al. [2017] propose a targeted minimum loss estimation (TMLE) approach that refines given nuisance estimators so that they asymptotically satisfy (2.4) and (2.5). The TMLE-based approach iteratively updates the initial nuisance estimators μ_n and α_n using a reweighted generalized linear regression algorithm with offset to produce *targeted* estimators μ^*_n and α^*_n satisfying the empirical score equations

$$\frac{1}{n}\sum_{i=1}^{n} s^*_n(A_i,W_i)\{Y_i - \mu^*_n(A_i,W_i)\} = o(n^{-\frac{1}{2}}); \tag{2.17}$$

$$\frac{1}{n}\sum_{i=1}^{n} r^*_n(A_i,W_i)\alpha^*_n(A_i,W_i) - \psi_n(r^*_n) = o(n^{-\frac{1}{2}}). \tag{2.18}$$

Here, s^*_n and r^*_n estimate the data-dependent nuisance functions $s^*_{n,0} := \Pi_{\mu^*_n}(\alpha_0 - \alpha^*_n)$ and $r^*_{n,0} := \Pi_{\alpha^*_n}(\mu_0 - \mu^*_n)$. They can be obtained, for example, by univariate empirical risk minimization with features $\mu^*_n(A_1,W_1),\ldots,\mu^*_n(A_n,W_n)$ and $\alpha^*_n(A_1,W_1),\ldots,\alpha^*_n(A_n,W_n)$, respectively; we provide details in Appendix 2.D.5. Under suitable regularity conditions, (2.17) and (2.18) imply that (2.4) and (2.5) hold asymptotically, so the one-step debiased estimator based on μ^*_n and α^*_n is indeed DRAL.

As discussed in the Introduction, the TMLE-based procedure for doubly robust inference can be computationally intensive because it requires estimating two univariate nuisance functions at each iteration, including $r^*_{n,0}$ and $s^*_{n,0}$ in the final step. In addition, the final estimation of $r^*_{n,0}$ and $s^*_{n,0}$ may introduce unnecessary finite-sample bias in the one-step debiased estimator τ_n, since the empirical moment equations (2.4) and (2.5) are then solved only approximately. We also note that neither van der Laan [2014] nor Benkeser et al. [2017] formally show that their iterative refinement procedures preserve the convergence properties of the original nuisance estimators while enforcing the empirical orthogonality equations. Such a result would be needed to ensure that, under consistent estimation of both nuisance

functions, the refinement step does not degrade the resulting estimator or inference. In the following section, we propose a novel procedure for constructing one-step debiased estimators that avoids these issues and, unlike prior work, simultaneously applies to a whole class of target parameters.

2.E Estimator irregularity under inconsistent nuisance estimation

As emphasized in the main text, Theorem 2.4 shows that the isocalibrated DML estimator is DRAL, and is regular and efficient when both nuisance functions are consistently estimated, even if one converges at an insufficient rate. Recall that an asymptotically linear estimator $\widehat{\psi}_n$ of ψ_0 with influence function f_0 is *regular* at P_0 for the functional $\Psi : \mathcal{M} \to \mathbb{R}$ if, for every one-dimensional regular parametric submodel $\{P_{0,t} : t \in (-\epsilon, \epsilon)\} \subset \mathcal{M}$ through P_0 at $t = 0$,

$$n^{1/2}\{\widehat{\psi}_n - \Psi(P_{0,t_n})\} \xrightarrow{d} N(0, P_0 f_0^2)$$

under sampling from P_{0,t_n}, where $t_n := tn^{-1/2}$ and $t \in \mathbb{R}$. Regularity means that inference for ψ_0 remains stable under local perturbations of the data-generating distribution.

When one nuisance function is inconsistently estimated, the calibrated estimator is typically irregular under a nonparametric model: it remains asymptotically linear, but with an influence function that differs from the nonparametric efficient influence function for the target parameter [Van der Vaart, 2000, Dukes et al., 2024]. This distinction matters because irregular estimators can have poor finite-sample behavior; in particular, there may be nearby distributions at which very large samples are needed before the asymptotic approximation is reliable [Leeb and Pötscher, 2005]. Thus, regularity is desirable, and in practice one should still aim to estimate both nuisance functions as accurately as possible.

At the same time, this irregularity is not unique to calibrated DML. Standard one-step debiased estimators, when one nuisance is inconsistently estimated, are typically neither regular nor root-n consistent, let alone asymptotically linear [Dukes et al., 2024]. Calibrated DML can nevertheless provide *pointwise*, rather than uniform, root-n inference in such settings. Thus, at laws where standard one-step estimators may fail to achieve root-n convergence, the isocalibrated DML estimator can remain asymptotically linear and support pointwise inference.

The following theorem characterizes the local asymptotic behavior of the isocalibrated

DML estimator in the irregular case where one nuisance function is inconsistently estimated; its proof is given in Appendix A.5.3. It shows that the estimator is regular and efficient for a projection-based oracle parameter Ψ_0, which agrees with the original target parameter $P \mapsto \Psi(P) := \psi_P(\mu_P)$ at P_0. Consequently, the isocalibrated DML procedure yields regular and locally uniformly valid inference for this projection estimand, even under least favorable local perturbations of P_0. Although the projection estimand is typically not of direct interest, this characterization lets us quantify the local asymptotic bias for inference on Ψ under such perturbations.

At a high level, when one nuisance is inconsistent, the oracle parameter replaces it by its best one-dimensional projection onto the model generated by the consistently estimated nuisance. To define this oracle parameter, we first introduce additional notation. Define the affine spaces

$$\overline{\Theta}_{\alpha_0,P} := \{\overline{\mu}_0 + f \circ \alpha_0 : f : \mathbb{R} \to \mathbb{R}\} \cap L^2(P), \qquad \overline{\Theta}_{\mu_0,P} := \{\overline{\alpha}_0 + g \circ \mu_0 : g : \mathbb{R} \to \mathbb{R}\} \cap L^2(P),$$

where f and g range over all one-dimensional real-valued transformations. Let the corresponding $L^2(P)$-orthogonal projections be

$$\overline{\Pi}_{\alpha_0,P}\mu_P := \operatorname*{argmin}_{\mu \in \overline{\Theta}_{\alpha_0,P}} \|\mu_P - \mu\|_P, \qquad \overline{\Pi}_{\mu_0,P}\alpha_P := \operatorname*{argmin}_{\alpha \in \overline{\Theta}_{\mu_0,P}} \|\alpha_P - \alpha\|_P.$$

We then define the projection-based oracle parameter Ψ_0, with the definition depending on which nuisance function is estimated inconsistently, by

$$P \mapsto \Psi_0(P) := \begin{cases} \overline{\psi}_{\mu_0,P}(\mu_P) & \text{if } \overline{\mu}_0 = \mu_0, \\ \psi_P(\overline{\Pi}_{\alpha_0,P}\mu_P) & \text{if } \overline{\alpha}_0 = \alpha_0. \end{cases}$$

Here,

$$\overline{\psi}_{\mu_0,P} : h \mapsto \langle \overline{\Pi}_{\mu_0,P}\alpha_P, h\rangle_P$$

is an oracle approximation of the linear functional $\psi_P : h \mapsto \langle \alpha_P, h\rangle_P$.

The orthogonality conditions defining the projections $\overline{\Pi}_{\mu_0,P_0}\alpha_0$ and $\overline{\Pi}_{\alpha_0,P_0}\mu_0$ imply that $\Psi_0(P_0) = \Psi(P_0)$, so the oracle and target parameters coincide at the true distribution P_0. More generally, when $\overline{\mu}_0 = \mu_0$, the parameters Ψ_0 and Ψ agree on the submodel

$$\Big\{P \in \mathcal{M} : \mu_P = g \circ \mu_0 \text{ for some } g : \mathbb{R} \to \mathbb{R}\Big\},$$

that is, on all laws P whose outcome regression is a one-dimensional transformation of μ_0.

This follows because $\overline{\psi}_{\mu_0,P}$ has Riesz representer $\overline{\Pi}_{\mu_0,P}\alpha_P$, which is the Riesz representer of the restriction of ψ_P to functions of the form $g \circ \mu_0$. Similarly, when $\overline{\alpha}_0 = \alpha_0$, the parameters Ψ_0 and Ψ agree on the submodel

$$\Big\{P \in \mathcal{M} : \alpha_P = f \circ \alpha_0 \text{ for some } f : \mathbb{R} \to \mathbb{R}\Big\},$$

that is, on all laws P whose Riesz representer is a one-dimensional transformation of α_0.

In the case where μ_0 is consistently estimated ($\overline{\mu}_0 = \mu_0$), Ψ_0 reduces to the parameter $P \mapsto \overline{\psi}_{\mu_0,P}(\mu_P)$, which uses the outcome regression $\mu_0(A, W)$ as a dimension reduction of (A, W). Alternatively, in the case where α_0 is consistently estimated ($\overline{\alpha}_0 = \alpha_0$), Ψ_0 reduces to the parameter $P \mapsto \psi_P(\overline{\Pi}_{\alpha_0,P}\mu_P)$, which uses the Riesz representer $\alpha_0(A, W)$ as a dimension reduction of (A, W).

Theorem 2.13 (Behavior under local perturbations)**.** *If the conditions of Theorem 2.4 hold and either $\overline{\mu}_0 = \mu_0$ or $\overline{\alpha}_0 = \alpha_0$, then χ_0 is the efficient influence function of Ψ_0 at P_0. Consequently, τ_n^* is regular, asymptotically linear, and efficient at P_0 for the oracle projection parameter Ψ_0, with influence function χ_0. In particular, under sampling from any local perturbation P_{0,t_n} of P_0,*

$$n^{1/2}\{\tau_n^* - \Psi_0(P_{0,t_n})\}$$

converges in distribution to a mean-zero normal random variable with variance σ_0^2, as defined in Theorem 2.4.

Theorem 2.13 shows that, even under sampling from least-favorable local perturbations, the isocalibrated DML estimator remains regular and yields locally uniformly valid inference for the projection-based oracle parameter Ψ_0. Although Ψ_0 is not itself of primary interest, it coincides with the target estimand $\Psi(P_0) = \psi_0(\mu_0)$ at P_0. Thus, even when inference for Ψ is irregular, the isocalibrated DML estimator remains regular for a well-defined, closely related parameter. As a special case, Theorem 2.13 implies that τ_n^* is regular and asymptotically linear for the true target parameter Ψ over all local perturbations that leave the consistently estimated nuisance function unchanged, in agreement with Theorem 2 of Dukes et al. [2024]. More generally, it is regular over all local perturbations P_{0,t_n} for which $\mu_{P_{0,t_n}}$ remains in $\overline{\Theta}_{\mu_0,P_{0,t_n}}$ when $\overline{\mu}_0 = \mu_0$, or for which $\alpha_{P_{0,t_n}}$ remains in $\overline{\Theta}_{\alpha_0,P_{0,t_n}}$ when $\overline{\alpha}_0 = \alpha_0$, since Ψ_0 and

Ψ agree along such perturbations.

The following corollary characterizes the asymptotic bias of the calibrated DML estimator for Ψ under local perturbations, and is an immediate consequence of Theorem 7 of van der Laan et al. [2023].

Corollary 2.14 (Asymptotic bias under local perturbations)**.** *Assume the setup of Theorem 2.13. Under sampling from any local perturbation P_{0,t_n} of P_0 with unit score $s \in L_0^2(P_0)$ $\left(\|s\|_{L^2(P_0)} = 1\right)$ and perturbation $t_n := tn^{-1/2}$ for $t \in \mathbb{R}$,*

$$n^{1/2}\{\tau_n^* - \Psi(P_{0,t_n})\} \xrightarrow{d} N\left(b_0(t),\, \sigma_0^2\right),$$

where $b_0(t) := t\langle s, \chi_0 - D_0\rangle$. Moreover, $|b_0(t)| \leq |t|\,\|\chi_0 - D_0\|$.

Consequently, the local asymptotic bias under P_{0,t_n} equals $b_0(t)$ and, in the worst case, is bounded by $|t|\,\|\chi_0 - D_0\|$. In particular, this bound can be controlled in terms of misspecification of the nuisance limits; for example,

$$\|\chi_0 - D_0\| \lesssim \|\overline{\mu}_0 - \mu_0\|_P \vee \|\overline{\alpha}_0 - \alpha_0\|,$$

so the local bias is small when the nuisance estimators are only mildly misspecified. Thus, even under local perturbations, mild inconsistency of the nuisance functions may lead to only a small local asymptotic bias for inference on Ψ.

2.F Extensions of calibrated DML

2.F.1 Extension to mixed bias parameters

In this section, we provide a brief overview of how our doubly robust inference framework can be extended to a general class of parameters with the mixed bias property, as introduced in Rotnitzky et al. [2021]. While our primary focus in this paper remains on the class of linear functionals discussed in Section 2.2.1, it is important to emphasize that our technique and results can be readily generalized to parameters exhibiting the mixed bias property.

We focus on debiasing mixed bias remainders of the following form using isotonic calibration:

$$\int w(z)\{\alpha_n(w,a) - \alpha_0(w,a)\}\{\gamma_n(w,a) - \gamma_0(w,a)\}P_0(dz)$$

where $z \mapsto w(z)$ is a known weighting function, and $\alpha_0 := \text{argmin}_\alpha E_0[w(Z)\{\alpha(A,W)\}^2 - 2m_1(A,W,\alpha)]$ and $\gamma_0 := \text{argmin}_\gamma E_0[w(Z)\{\gamma(A,W)\}^2 - 2m_2(A,W,\gamma)]$ are weighted Riesz representers for the linear functionals $\alpha \mapsto E_0[m_1(A,W,\alpha)]$ and $\gamma \mapsto E_0[m_2(A,W,\gamma)]$. Mixed bias parameters, as defined in Rotnitzky et al. [2021], have remainders of this form. More generally, parameters may involve second-order remainders consisting of multiple such mixed bias terms, each of which can be linearized separately using the results of this section.

Given estimators α_n and γ_n of α_0 and γ_0, we propose using the isocalibrated nuisance estimators $\alpha_n^* = g_{n,1} \circ \alpha_n$ and $\gamma_n^* = g_{n,2} \circ \gamma_n$, where

$$\begin{aligned} g_{n,1} &\in \underset{g \in \mathcal{F}_{\text{iso}}}{\text{argmin}} \sum_{i=1}^n w(Z_i)\big\{g(\alpha_n(A_i,W_i))\big\}^2 - 2m_1(Z_i, g \circ \alpha_n), \\ g_{n,2} &\in \underset{g \in \mathcal{F}_{\text{iso}}}{\text{argmin}} \sum_{i=1}^n w(Z_i)\big\{g(\gamma_n(A_i,W_i))\big\}^2 - 2m_2(Z_i, g \circ \gamma_n). \end{aligned} \tag{2.19}$$

Notably, the first-order conditions of the minimizing solutions guarantee that

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^n [w(Z_i)(g \circ \alpha_n^*)(W_i,A_i)\alpha_n^*(W_i,A_i) - m_1(W_i,A_i,g \circ \alpha_n^*)] = 0, \\ \frac{1}{n}\sum_{i=1}^n [w(Z_i)(g \circ \gamma_n^*)(W_i,A_i)\gamma_n^*(W_i,A_i) - m_2(W_i,A_i,g \circ \gamma_n^*)] = 0 \end{aligned}$$

for all transformations $g : \mathbb{R} \to \mathbb{R}$. As in Algorithms 1 and 2, the initial nuisance estimators can be cross-fitted, and confidence intervals can be constructed using the bootstrap. Our proof techniques can be readily extended to establish that isotonic calibration guarantees doubly robust asymptotic linearity of mixed bias remainders.

Example 2.15 (Calibrated DML for expected conditional covariance)**.** Let π_n and m_n be estimators of $w \mapsto E_0[A \mid W = w]$ and $w \mapsto E_0[Y \mid W = w]$, respectively. The expected conditional covariance parameter

$$E_0\left[\{A - \pi_0(W)\}\{Y - m_0(W)\}\right] = E_0\left[\{A - \pi_0(W)\}^2\{\mu_0(1,W) - \mu_0(0,W)\}\right]$$

arises in the context of estimating the ATE under treatment effect homogeneity. Doubly robust inference for this parameter was studied in Dukes et al. [2024] and Bonvini et al. [2024]. This parameter is a linear functional of the outcome regression, but the functional depends on the nuisance function π_0. In this case, a modification of our procedure yields improved properties. A doubly robust estimator of the expected covariance parameter is

given by

$$\frac{1}{n}\sum_{i=1}^{n}\{A_i - \pi_n(W_i)\}\{Y_i - m_n(W_i)\}.$$

The second-order remainder of this estimator involves the cross-product term [Robinson, 1988]:

$$\langle \pi_n(W) - \pi_0(W), m_n - m_0(W)\rangle.$$

This mixed bias term is exactly of the form addressed in Section 2.2.2, involving an inner product between a regression function and a Riesz representer. The regression function m_0 corresponds to the least-squares loss $(z, m) \mapsto \{y - m(w)\}^2$, while the function π_0 is a Riesz representer of the linear functional $\alpha \mapsto E_0[A\alpha(W)]$, with Riesz loss corresponding to the usual least-squares loss $(z, \alpha) \mapsto \{a - \alpha(w)\}^2$.

Thus, a calibrated DML estimator of the expected covariance parameter is given by

$$\frac{1}{n}\sum_{i=1}^{n}\{A_i - \pi_n^*(W_i)\}\{Y_i - m_n^*(W_i)\},$$

where $\pi_n^* = g_n \circ \pi_n$ and $m_n^* = f_n \circ m_n$, with

$$g_n = \operatorname*{argmin}_{g\in\mathcal{F}_{\text{iso}}} \sum_{i=1}^{n} \{A_i - g(\pi_n(W_i))\}^2,$$
$$f_n = \operatorname*{argmin}_{f\in\mathcal{F}_{\text{iso}}} \sum_{i=1}^{n} \{Y_i - f(m_n(W_i))\}^2.$$

Doubly robust asymptotic linearity of this estimator follows from conditions analogous to those in Section 2.5.

2.F.2 A general template for higher-order debiasing

The debiasing strategy of Section 2.2.2 can be substantially generalized, as the analysis relies only on empirical orthogonality properties of the nuisance estimators and basic properties of projections. In this section, we present a general debiasing strategy that extends beyond this setting.

Our extension uses the following notation. Let $\Theta_{\mu,n,0} \subset \mathcal{H}$ be a linear subspace that contains the *regression error* $\mu_n - \mu_0$ (this space may depend on n), and let $\Theta_{\mu,n} \subset \Theta_{\mu,n,0}$ be a lower-dimensional *approximating* subspace. We view $\Theta_{\mu,n,0}$ as a working model for the directions in which μ_n may deviate from μ_0, and $\Theta_{\mu,n}$ as the tractable subset of those

directions on which we can enforce empirical orthogonality. Similarly, let $\Theta_{\alpha,n,0} \subseteq \mathcal{H}$ contain the *Riesz error* $\alpha_n - \alpha_0$, and let $\Theta_{\alpha,n} \subseteq \Theta_{\alpha,n,0}$ be an approximating subspace. We assume that the relevant affine spaces are closed under $\|\cdot\|$. We write $\Pi_{\mu,n,0}$, $\Pi_{\mu,n}$, $\Pi_{\alpha,n,0}$, and $\Pi_{\alpha,n}$ for the $\|\cdot\|$-orthogonal projections onto $\Theta_{\mu,n,0}$, $\Theta_{\mu,n}$, $\Theta_{\alpha,n,0}$, and $\Theta_{\alpha,n}$, respectively; for example, $\Pi_{\mu,n} f := \arg\min_{h \in \Theta_{\mu,n}} \|f - h\|$.

Two examples recover familiar constructions. First, the calibrated DML setup in Section 2.3.2 is obtained by taking $\Theta_{\mu,n} := \overline{\{f \circ \mu_n : f \colon \mathbb{R} \to \mathbb{R}\}}$ and $\Theta_{\mu,n,0} := \overline{\{g \circ (\mu_n, \mu_0) : g \colon \mathbb{R}^2 \to \mathbb{R}\}}$, with $\Theta_{\alpha,n}$ and $\Theta_{\alpha,n,0}$ defined analogously; in this case, $\Theta_{\mu,n}$ corresponds to calibration directions available through post hoc one-dimensional transformations of μ_n. Second, for a linear sieve given by a nested sequence $\mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \cdots \subseteq \mathcal{H}_\infty \subseteq \mathcal{H}$, one may take $\Theta_{\mu,n} := \mathcal{H}_{k(n)}$ and $\Theta_{\mu,n,0} := \mathcal{H}_\infty$, where $k(n) \uparrow \infty$ as the sample size increases; here $\Theta_{\mu,n}$ is the finite-dimensional space used at sample size n, while $\Theta_{\mu,n,0}$ is the limiting sieve space containing the regression error.

Our more general debiasing strategy constructs nuisance estimators μ_n and α_n that satisfy the empirical orthogonality (score) conditions

$$\frac{1}{n}\sum_{i=1}^{n} (\Pi_{\mu,n}\Delta_{\alpha,n})(A_i, W_i)\{Y_i - \mu_n(A_i, W_i)\} = 0, \tag{2.20}$$

$$\psi_n(\Pi_{\alpha,n}\Delta_{\mu,n}) - \frac{1}{n}\sum_{i=1}^{n} (\Pi_{\alpha,n}\Delta_{\mu,n})(A_i, W_i)\, \alpha_n(A_i, W_i) = 0. \tag{2.21}$$

In words, (2.20) requires that the residuals of μ_n are empirically orthogonal to the projection of the estimation error $\Delta_{\alpha,n}$ onto the approximating subspace $\Theta_{\mu,n}$ for $\mu_n - \mu_0$. If $\Theta_{\mu,n}$ happened to contain $\Delta_{\alpha,n}$ itself, then this condition reduces to enforcing

$$\frac{1}{n}\sum_{i=1}^{n} \Delta_{\alpha,n}(A_i, W_i)\{Y_i - \mu_n(A_i, W_i)\} \approx 0,$$

which sets the empirical estimate of the cross-product remainder $\langle \Delta_{\alpha,n}, \Delta_{\mu,n} \rangle$ to zero. If μ_n and α_n are sieve-type estimators that respectively minimize the least-squares and Riesz losses over $\Theta_{\mu,n}$ and $\Theta_{\alpha,n}$, then both conditions hold automatically by the corresponding first-order optimality conditions. Otherwise, given initial estimators $\mu_n^{(0)}$ and $\alpha_n^{(0)}$, these conditions can be enforced post hoc by defining $\mu_n := \mu_n^{(0)} + h_{n,\mu}$ and $\alpha_n := \alpha_n^{(0)} + h_{n,\alpha}$, where the adjustment terms are obtained via the offset sieve problems

$$h_{n,\mu} := \arg\min_{h \in \Theta_{\mu,n}} \sum_{i=1}^{n} \{Y_i - \mu_n^{(0)}(A_i, W_i) - h(A_i, W_i)\}^2,$$

$$h_{n,\alpha} := \arg\min_{h \in \Theta_{\alpha,n}} \sum_{i=1}^{n} h(A_i, W_i)^2 - 2\{\psi_n(h) - \langle \alpha_n^{(0)}, h \rangle_{P_n}\},$$

where $\Theta_{\mu,n}$ and $\Theta_{\alpha,n}$ are taken to be finite-dimensional and tuned to control overfitting. In high-dimensional settings, one can add ℓ^1 or ℓ^2 regularization, in which case the orthogonality equations typically hold only approximately, in view of the KKT conditions for the minimizers. The resulting regularization error must then be controlled, as in Bradic et al. [2019a] and Smucler et al. [2019].

In the following theorem, we denote $\widetilde{A}_{n,0} : z \mapsto \widetilde{s}_{n,0}(a,w)\{y - \mu_n(a,w)\}$, $\widetilde{B}_{n,0} := \psi_0(\widetilde{r}_{n,0}) + \phi_{\widetilde{r}_{n,0}} - \widetilde{r}_{n,0}\alpha_n$, and $\mathrm{Rem}_{\psi_n}(\widetilde{r}_{n,0}) := \psi_n(\widetilde{r}_{n,0}) - \psi_0(\widetilde{r}_{n,0}) - (P_n - P_0)\phi_{\widetilde{r}_{n,0}}$, where $\widetilde{s}_{n,0} := \Pi_{\mu,n}(\alpha_0 - \alpha_n)$ and $\widetilde{r}_{n,0} := \Pi_{\alpha,n}(\mu_0 - \mu_n)$.

Theorem 2.16 (Expansion of the cross-product remainder under general orthogonality)**.** *If μ_n satisfies (2.20), then we have the regression-based decomposition:*

$$\langle \alpha_0 - \alpha_n, \mu_n - \mu_0 \rangle = (P_n - P_0)\widetilde{A}_{n,0} - \langle (\Pi_{\mu,n,0} - \Pi_{\mu,n})\Delta_{\alpha,n},\, (\Pi_{\mu,n,0} - \Pi_{\mu,n})\Delta_{\mu,n} \rangle.$$

If α_n satisfies (2.21), then we have the Riesz-based decomposition:

$$\langle \alpha_0 - \alpha_n, \mu_n - \mu_0 \rangle = (P_n - P_0)\widetilde{B}_{n,0} - \langle (\Pi_{\alpha,n,0} - \Pi_{\alpha,n})\Delta_{\mu,n},\, (\Pi_{\alpha,n,0} - \Pi_{\alpha,n})\Delta_{\alpha,n} \rangle + \mathrm{Rem}_{\psi_n}(\widetilde{r}_{n,0}).$$

The decompositions in Theorem 2.16 are similar to those in Theorem 2.3. Each decomposition consists of an empirical process term, which is asymptotically linear under suitable conditions, and a higher-order cross-product remainder. In the regression-based case, these terms are $(P_n - P_0)\widetilde{A}_{n,0}$ and $-\big\langle (\Pi_{\mu,n,0} - \Pi_{\mu,n})\Delta_{\alpha,n},\, (\Pi_{\mu,n,0} - \Pi_{\mu,n})\Delta_{\mu,n} \big\rangle$, respectively. By the Cauchy–Schwarz inequality,

$$\big|\langle (\Pi_{\mu,n,0} - \Pi_{\mu,n})\Delta_{\alpha,n},\, (\Pi_{\mu,n,0} - \Pi_{\mu,n})\Delta_{\mu,n} \rangle\big| \le \|(\Pi_{\mu,n,0} - \Pi_{\mu,n})\Delta_{\alpha,n}\|\, \|\Delta_{\mu,n} - \Pi_{\mu,n}\Delta_{\mu,n}\|. \tag{2.22}$$

Consequently, the higher-order cross-product remainder is small if either the regression error $\Delta_{\mu,n}$ is well approximated by $\Theta_{\mu,n}$, or the projected Riesz error $\Pi_{\mu,n,0}\Delta_{\alpha,n}$ is well approximated by $\Theta_{\mu,n}$. In the special case $\Theta_{\mu,n,0} = \mathcal{H}$, this bound reduces to the product of approximation errors for the nuisance estimation errors, namely $\|\Delta_{\alpha,n} - \Pi_{\mu,n}\Delta_{\alpha,n}\|\, \|\Delta_{\mu,n} - \Pi_{\mu,n}\Delta_{\mu,n}\|$.

To provide intuition, we interpret the bound in (2.22) in the special case where $\Theta_{\mu,n} := \mathcal{H}_{k(n)}$ and $\Theta_{\mu,n,0} := \mathcal{H}_\infty$ for a linear sieve $\{\mathcal{H}_k\}_{k=1}^{\infty}$. In this setting, the right-hand side is

the product of two sieve approximation errors. The first term is small when the projected Riesz error $\Pi_{\mu,n,0}\Delta_{\alpha,n}$ is sufficiently smooth to be well approximated by the sieve space $\Theta_{\mu,n}$, while the second term is bounded by $\|\mu_n - \mu_0\|$ by standard projection properties. Theorem 2.16 can therefore be used to establish doubly robust asymptotic linearity of sieve-based autoDML estimators that use sieve estimators for both the Riesz representer and the outcome regression. As a special case, setting either nuisance estimator to the zero function recovers the classical result that plug-in sieve estimators are asymptotically linear without explicit debiasing.

Rather than working with the projected errors, a sufficient—though stronger than needed—condition for the higher-order cross-product remainders to be negligible is that the spaces $\Theta_{\mu,n}$ and $\Theta_{\alpha,n}$ provide good approximations to $\Delta_{\alpha,n}$ and $\Delta_{\mu,n}$, respectively. From a practical standpoint, however, directly choosing $\Theta_{\mu,n}$ to approximate $\Delta_{\alpha,n}$ is often unrealistic, since $\Delta_{\alpha,n}$ is typically high-dimensional and unknown. Calibrated DML avoids this curse of dimensionality by using μ_n and α_n as low-dimensional summaries and defining $\Theta_{\mu,n}$ and $\Theta_{\alpha,n}$ as classes of one-dimensional functions composed with them.

2.G Background on calibration

In this section, we provide background on calibration using binning and isotonic regression. For additional details, we refer to van der Laan and Alaa [2025] for a general treatment of histogram binning and isotonic calibration; see also Gupta et al. [2020], Gupta and Ramdas [2021] for histogram binning, Van Der Laan et al. [2023] and van der Laan et al. [2024b] for applications of isotonic calibration in causal inference, and Whitehouse et al. [2024] for calibration with respect to general loss functions.

2.G.1 Empirical calibration via histogram binning

Let $(z, \nu) \mapsto \ell(z, \nu)$ be a loss function for some nuisance function $\nu_0 \in \mathcal{H}$. An estimator ν_n is said to be *ℓ–empirically calibrated* for ν_0 if it satisfies the self-consistency property

$$\frac{1}{n}\sum_{i=1}^{n} \ell(Z_i, \nu_n) = \min_{\theta} \frac{1}{n}\sum_{i=1}^{n} \ell(Z_i, \theta \circ \nu_n)\,, \tag{2.23}$$

where the minimum is taken over all real-valued functions $\theta : \mathbb{R} \to \mathbb{R}$. That is, the empirical risk of ν_n cannot be improved through transformation by any one-dimensional mapping.

When the loss ℓ is sufficiently smooth, the first-order equations characterizing the empirical risk minimization problem imply that ℓ–empirical calibration is equivalent to requiring that, for each map $\theta : \mathbb{R} \to \mathbb{R}$,

$$\frac{1}{n}\sum_{i=1}^{n} \frac{\partial}{\partial \varepsilon} \ell(Z_i, \nu_n + \varepsilon\theta \circ \nu_n)\Big|_{\varepsilon=0} = 0\,. \tag{2.24}$$

A simple, distribution-free method for calibrating nuisance estimators is *histogram binning*, such as uniform-mass (quantile) binning [Gupta et al., 2020, Gupta and Ramdas, 2021]. Given an estimated function ν_n, the calibration dataset $\mathcal{C}_n = \{Z_i\}_{i=1}^n$, and a loss function $\ell(z, \nu)$, the range of predictions $\nu_n(Z_i)$ is partitioned into K disjoint bins $\{B_k\}_{k=1}^K$ in an outcome-independent manner—for example, using quantiles of $\{\nu_n(Z_i)\}_{i=1}^n$. The calibrated estimator $\nu_n^* := \theta_n \circ \nu_n$ is then obtained by minimizing the empirical loss within each bin:

$$\theta_n(t) := \operatorname*{argmin}_{c \in \mathbb{R}} \sum_{i=1}^{n} \mathbb{1}\{\nu_n(Z_i) \in B_{k(t)}\}\, \ell(Z_i, c),$$

where $k(t)$ denotes the index of the bin containing $t \in \nu_n(\mathcal{Z})$. For least-squares loss, this reduces to computing the average of the observed outcomes within each bin. Since any transformation of a predictor preserves its piecewise constant structure, the empirical risk cannot be reduced by further transforming its predictions, and hence it is *empirically ℓ-calibrated*.

While histogram binning guarantees empirical calibration, it does not ensure good out-of-sample performance. With a fine partition ($K = n$), the estimator may interpolate the data and overfit, leading to poor generalization. Conversely, with a coarse partition ($K = 1$), the estimator is well-calibrated but lacks predictive power, collapsing to a constant predictor. This trade-off highlights the importance of choosing K carefully: too few bins limit expressiveness, while too many increase variance and degrade calibration. Under appropriate choices of K, histogram binning yields asymptotically calibrated estimators, with the conditional ℓ^2-calibration error satisfying $\mathrm{Cal}_{\ell^2}(\nu_n^*) = O_p\left(\frac{K \log(n/K)}{n}\right)$ [Whitehouse et al., 2024, van der Laan and Alaa, 2025]. Thus, tuning K, for example via cross-validation, is essential to balance calibration and accuracy. In the next section, we adopt a data-adaptive alternative—*isotonic calibration*—which flexibly enforces monotonicity while maintaining

distribution-free guarantees [Zadrozny and Elkan, 2001, Niculescu-Mizil and Caruana, 2005b, Van Der Laan et al., 2023].

2.G.2 Empirical calibration via isotonic regression

Isotonic calibration [Zadrozny and Elkan, 2002, Niculescu-Mizil and Caruana, 2005a] is a data-adaptive histogram binning method that learns the bins using isotonic regression, a nonparametric technique traditionally used for estimating monotone functions [Barlow and Brunk, 1972, Groeneboom and Lopuhaa, 1993]. Specifically, the bins are determined by minimizing an empirical mean least-squares criterion under the constraint that the calibrated predictor is a non-decreasing monotone transformation of the original predictor. Isotonic calibration is motivated by the heuristic that, for a good predictor $z \mapsto \nu_n(z)$, the optimal transformation of the predictor should be approximately monotone as a function of ν_n. For example, for a perfect predictor, the optimal transformation should be the identity function. Even when this map is not approximately monotone, isotonic regression ensures that the quality of the initial predictor is not harmed, since the identity transformation—being monotone—can always be learned. Isotonic calibration is distribution-free—it does not rely on monotonicity assumptions—and, unlike traditional histogram binning, it is tuning-parameter-free and naturally preserves the mean least-squares of the original predictor (as the identity transformation is monotone) [Van Der Laan et al., 2023].

For clarity, we focus on the regression case where ℓ denotes the least-squares loss. Formally, isotonic calibration takes a fitted nuisance estimator ν_n and a calibration dataset $\mathcal{C}_n$ to produce a calibrated model $\nu_n^* := \theta_n \circ \nu_n$, where $\theta_n : \mathbb{R} \to \mathbb{R}$ is an isotonic step function obtained by solving the optimization problem:

$$\theta_n \in \underset{\theta \in \mathcal{F}_{\text{iso}}}{\operatorname{argmin}} \sum_{i=1}^{n} \{Y_i - \theta(\nu_n(W_i, A_i))\}^2 , \tag{2.25}$$

where $\mathcal{F}_{\text{iso}}$ denotes the set of all univariate, piecewise constant functions that are monotonically nondecreasing. Following Groeneboom and Lopuhaa [1993], we consider the unique càdlàg piecewise constant solution to this problem, which has jumps only at observed values $\{\nu_n(W_i, A_i) : i \in [n]\}$. The first-order optimality conditions of this convex problem imply that the isotonic solution θ_n acts as a binning calibrator over a data-adaptive set of bins determined by the jump points of the step function. Thus, isotonic calibration

yields perfect empirical calibration. Specifically, for any transformation $g : \mathbb{R} \to \mathbb{R}$, the perturbed step function $\varepsilon \mapsto \theta_n + \varepsilon(g \circ \theta_n)$ remains isotonic for all sufficiently small ε such that $|\varepsilon| \sup_{t \in \nu_n(\mathcal{Z})} |(g \circ \theta_n)(t)|$ is less than the maximum jump size of θ_n, given by $\sup_{t \in \nu_n(\mathcal{Z})} |\theta_n(t) - \theta_n(t-)|$. Since θ_n minimizes the empirical mean least-squares over all isotonic functions, it follows that for each $g : \mathbb{R} \to \mathbb{R}$, the following condition holds:

$$\frac{d}{d\varepsilon} \frac{1}{2} \sum_{i=1}^{n} \{Y_i - \theta_n(\nu_n(W_i, A_i)) - \varepsilon g(\theta_n(\nu_n(W_i, A_i)))\}^2 \Big|_{\varepsilon=0} = \sum_{i=1}^{n} g(\nu_n^*(Z_i))\{Y_i - \nu_n^*(Z_i)\} = 0.$$

These orthogonality conditions are equivalent to perfect empirical calibration. In particular, by taking $g(t) = 1\{t = \theta_n(\nu_n(z))\}$, we conclude that the isotonic calibrated predictor ν_n^* is empirical calibrated.

2.G.3 Benefits and trade-offs in empirical calibration

In the context of debiased machine learning, empirical calibration is valuable not only for enabling doubly robust inference but also for improving estimator stability and predictive accuracy. In our framework, empirical calibration is defined relative to a loss function $\ell(z, \nu)$ for a nuisance function $\nu \in \mathcal{H}$, and requires that ν_n satisfies the self-consistency condition

$$\frac{1}{n} \sum_{i=1}^{n} \ell(Z_i, \nu_n) = \min_{\theta} \frac{1}{n} \sum_{i=1}^{n} \ell(Z_i, \theta \circ \nu_n),$$

where the minimum is taken over all univariate transformations $\theta : \mathbb{R} \to \mathbb{R}$. This condition ensures that the predictions ν_n are optimal with respect to empirical risk and cannot be improved by recalibration.

In the case of least-squares loss $\ell(z, \mu) = \{y - \mu(a, w)\}^2$, calibration requires that $\mu_n(a, w)$ equals the average outcome among individuals with the same predicted value. That is,

$$\mu_n(a, w) = \frac{\sum_{i=1}^{n} 1\{\mu_n(A_i, W_i) = \mu_n(a, w)\} Y_i}{\sum_{i=1}^{n} 1\{\mu_n(A_i, W_i) = \mu_n(a, w)\}},$$

ensuring that μ_n does not systematically over- or under-predict outcomes. For the Riesz loss $\ell(z, \alpha) = \alpha^2(a, w) - 2\psi_n(\alpha)$, associated with estimating a counterfactual mean $\psi_0 = E_0\{\mu_0(1, W)\}$, calibration of the Riesz representer α_n implies approximate covariate balance across estimated inverse propensity strata. Specifically, the calibration condition

$$\frac{1}{n} \sum_{i=1}^{n} A_i \alpha_n(1, W_i) f(\alpha_n(1, W_i)) = \frac{1}{n} \sum_{i=1}^{n} f(\alpha_n(1, W_i))$$

must hold for all functions f, ensuring that higher weights contribute meaningfully to bal-

ancing, rather than inflating variance [Deshpande and Kuleshov, 2023, van der Laan et al., 2024b].

To ensure good performance of the empirically calibrated estimator, it is important that calibration holds not only on the observed data but also at the population level. Notably, perfect empirical calibration can be achieved trivially by overfitting the labels—for example, by interpolating the outcomes—so it is crucial to use calibration procedures that generalize well. Formally, a function ν is said to be perfectly ℓ-*calibrated* with respect to a loss function $\ell(z, \nu)$ if it minimizes the expected loss over all one-dimensional transformations of itself:

$$E_P[\ell(Z, \nu)] = \inf_{\theta} E_P[\ell(Z, \theta \circ \nu)],$$

where the infimum is taken over all functions $\theta : \mathbb{R} \to \mathbb{R}$. The notion of ℓ-calibration generalizes common forms of calibration, such as regression and quantile calibration [Roth, 2022], which correspond to the least-squares and pinball (hinge) losses, respectively [Deng et al., 2023]. Related ideas have also been explored in the context of predictive modeling [Whitehouse et al., 2024].

2.H Implementation of calibrated DML

An `R` package `calibratedDML` as well as `Python` code implementing CDML can be found on GitHub at the following link: https://github.com/Larsvanderlaan/calibratedDML.

The following subsections provide `R` and `Python` code for calibrating inverse propensity weights and outcome regressions. These calibrated models can be directly used within DML frameworks to construct isocalibrated DML estimators of the average treatment effect. Here, we implement isotonic regression using `xgboost`, allowing control over the maximum tree depth and the minimum number of observations in each constant segment of the isotonic regression fit.

2.H.1 R Code

```
# Function: isoreg_with_xgboost
# Purpose: Fits isotonic regression using XGBoost.
# Inputs:
#   - x: A vector or matrix of predictor variables.
#   - y: A vector of response variables.
```

```
#   - max_depth: Maximum depth of the trees in XGBoost (default = 15).
#   - min_child_weight: Minimum sum of instance weights (Hessian)
#       needed in a child node (default = 20).
# Returns:
#   - A function that takes a new predictor variable x
#     and returns the model's predicted values.

isoreg_with_xgboost <- function(x, y, max_depth = 15, min_child_weight = 20) {

  # Create an XGBoost DMatrix object from the data
  data <- xgboost::xgb.DMatrix(data = as.matrix(x), label = as.vector(y))

  # Set parameters for the monotonic XGBoost model
  params <- list(max_depth = max_depth,
                 min_child_weight = min_child_weight,
                 monotone_constraints = 1,  # Enforce monotonic increase
                 eta = 1, gamma = 0,
                 lambda = 0)

  # Train the model with one boosting round
  iso_fit <- xgboost::xgb.train(params = params,
                                data = data,
                                nrounds = 1)

  # Prediction function for new data
  fun <- function(x) {
    data_pred <- xgboost::xgb.DMatrix(data = as.matrix(x))
    pred <- predict(iso_fit, data_pred)
    return(pred)
  }
  return(fun)
}

# Function: calibrate_inverse_weights
# Purpose: Calibrates inverse weights using isotonic regression
#          with XGBoost for two propensity scores.
# Inputs:
#   - A: Binary indicator variable.
#   - pi1: Estimated propensity score for treatment group A = 1.
#   - pi0: Estimated propensity score for control group A = 0.
# Returns:
#   - A list containing calibrated inverse weights for each group:
```

```
#       - alpha1_star: Calibrated inverse weights for A = 1.
#       - alpha0_star: Calibrated inverse weights for A = 0.

calibrate_inverse_weights <- function(A, pi1, pi0) {

  # Calibrate pi1 using monotonic XGBoost
  calibrator_pi1 <- isoreg_with_xgboost(pi1, A)
  pi1_star <- calibrator_pi1(pi1)

  # Set minimum truncation level for treated group
  c1 <- min(pi1_star[A == 1])
  pi1_star = pmax(pi1_star, c1)
  alpha1_star <- 1 / pi1_star

  # Calibrate pi0 using monotonic XGBoost
  calibrator_pi0 <- isoreg_with_xgboost(pi0, 1 - A)
  pi0_star <- calibrator_pi0(pi0)

  # Set minimum truncation level for control group
  c0 <- min(pi0_star[A == 0])
  pi0_star = pmax(pi0_star, c0)
  alpha0_star <- 1 / pi0_star

  # Return calibrated inverse weights for both groups
  return(list(alpha1_star = alpha1_star, alpha0_star = alpha0_star))
}

# Function: calibrate_outcome_regression
# Purpose: Calibrates outcome regression predictions using isotonic
#          regression with XGBoost.
# Inputs:
#   - Y: Observed outcomes.
#   - mu1: Predicted outcome for treated group.
#   - mu0: Predicted outcome for control group.
#   - A: Binary indicator variable.
# Returns:
#   - A list containing calibrated predictions for each group:
#       - mu1_star: Calibrated predictions for A = 1.
#       - mu0_star: Calibrated predictions for A = 0.

calibrate_outcome_regression <- function(Y, mu1, mu0, A) {
```

```
  # Calibrate mu1 using monotonic XGBoost for treated group
  calibrator_mu1 <- isoreg_with_xgboost(mu1[A==1], Y[A==1])
  mu1_star <- calibrator_mu1(mu1)

  # Calibrate mu0 using monotonic XGBoost for control group
  calibrator_mu0 <- isoreg_with_xgboost(mu0[A==0], Y[A==0])
  mu0_star <- calibrator_mu0(mu0)

  # Return calibrated values for both groups
  return(list(mu1_star = mu1_star, mu0_star = mu0_star))
}
```

2.H.2 Python code

```
import xgboost as xgb
import numpy as np

def isoreg_with_xgboost(x, y, max_depth=15, min_child_weight=20):
    """
    Fits isotonic regression using XGBoost with monotonic constraints to ensure
    non-decreasing predictions as the predictor variable increases.

    Args:
        x (np.array): A vector or matrix of predictor variables.
        y (np.array): A vector of response variables.
        max_depth (int, optional): Maximum depth of the trees in XGBoost.
                                   Default is 15.
        min_child_weight (float, optional): Minimum sum of instance weights
                                            needed in a child node. Default is 20.

    Returns:
        function: A prediction function that takes a new predictor variable x
                  and returns the model's predicted values.

    Example:
        >>> x = np.array([[1], [2], [3]])
        >>> y = np.array([1, 2, 3])
        >>> model = isoreg_with_xgboost(x, y)
        >>> model(np.array([[1.5], [2.5]]))
    """
```

```
    # Create an XGBoost DMatrix object from the data
    data = xgb.DMatrix(data=np.asarray(x), label=np.asarray(y))

    # Set parameters for the monotonic XGBoost model
    params = {
        'max_depth': max_depth,
        'min_child_weight': min_child_weight,
        'monotone_constraints': "(1)",  # Enforce monotonic increase
        'eta': 1,
        'gamma': 0,
        'lambda': 0
    }

    # Train the model with one boosting round
    iso_fit = xgb.train(params=params, dtrain=data, num_boost_round=1)

    # Prediction function for new data
    def predict_fn(x):
        """
        Predicts output for new input data using the trained isotonic regression model.

        Args:
            x (np.array): New predictor variables as a vector or matrix.

        Returns:
            np.array: Predicted values.
        """
        data_pred = xgb.DMatrix(data=np.asarray(x))
        pred = iso_fit.predict(data_pred)
        return pred

    return predict_fn


def calibrate_inverse_weights(A, pi1, pi0):
    """
    Calibrates inverse weights using isotonic regression with XGBoost for two propensity scores.

    Args:
        A (np.array): Binary indicator variable.
        pi1 (np.array): Propensity score for treatment group (A = 1).
```

```
        pi0 (np.array): Propensity score for control group (A = 0).

    Returns:
        dict: Contains calibrated inverse weights:
              - alpha1_star: Inverse weights for A = 1.
              - alpha0_star: Inverse weights for A = 0.
    """
    calibrator_pi1 = isoreg_with_xgboost(pi1, A)
    pi1_star = calibrator_pi1(pi1)
    c1 = np.min(pi1_star[A == 1])
    pi1_star = np.maximum(pi1_star, c1)
    alpha1_star = 1 / pi1_star

    calibrator_pi0 = isoreg_with_xgboost(pi0, 1 - A)
    pi0_star = calibrator_pi0(pi0)
    c0 = np.min(pi0_star[A == 0])
    pi0_star = np.maximum(pi0_star, c0)
    alpha0_star = 1 / pi0_star

    return {'alpha1_star': alpha1_star, 'alpha0_star': alpha0_star}

def calibrate_outcome_regression(Y, mu1, mu0, A):
    """
    Calibrates outcome regression predictions using isotonic regression with XGBoost.

    Args:
        Y (np.array): Observed outcomes.
        mu1 (np.array): Predicted outcome for the treated group (A = 1).
        mu0 (np.array): Predicted outcome for the control group (A = 0).
        A (np.array): Binary treatment indicator (1 for treatment, 0 for control).

    Returns:
        dict: Calibrated predictions for each group:
              - 'mu1_star': Calibrated predictions for A = 1.
              - 'mu0_star': Calibrated predictions for A = 0.
    """
    # Calibrate mu1 using treated group data (A = 1)
    calibrator_mu1 = isoreg_with_xgboost(mu1[A == 1], Y[A == 1])
    mu1_star = calibrator_mu1(mu1)  # Apply calibrator to all mu1 values

    # Calibrate mu0 using control group data (A = 0)
    calibrator_mu0 = isoreg_with_xgboost(mu0[A == 0], Y[A == 0])
```

```
    mu0_star = calibrator_mu0(mu0)  # Apply calibrator to all mu0 values

    return {'mu1_star': mu1_star, 'mu0_star': mu0_star}
```

2.I Additional details on numerical experiments

2.I.1 Complete results for experiment 1

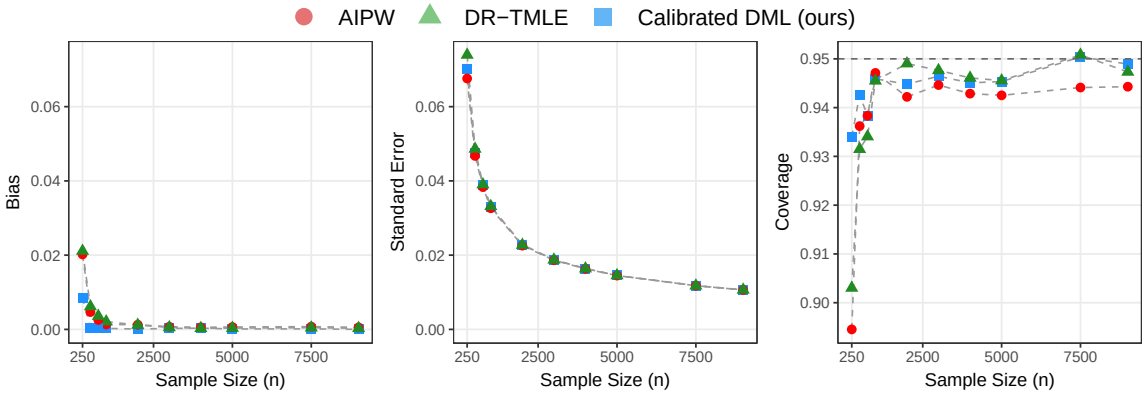


Figure 2.4: Empirical relative bias, standard error, and 95% confidence interval coverage for isocalibrated DML (IC-DML), DR-TMLE and AIPW estimators under both doubly consistent estimation of the outcome regression and treatment mechanism. Both propensity score and outcome regression are estimated consistently.

2.I.2 Complete results for experiment 2

Table 2.4: Evaluation of absolute bias, scaled root mean square error (RMSE), and 95% confidence-interval coverage. Both outcome regression and propensity scores are estimated with gradient-boosted trees.

Dataset	Bias		RMSE		Coverage	
	AIPW	C-DML	AIPW	C-DML	AIPW	C-DML
ACIC-2017 (17)	0.031	**0.026**	0.095	**0.094**	**0.95**	0.94
ACIC-2017 (18)	**0.21**	0.28	0.7	**0.58**	0.27	**0.64**
ACIC-2017 (19)	0.082	**0.057**	0.43	0.43	0.96	0.96
ACIC-2017 (20)	1.6	**0.2**	2	**1.4**	0.32	**0.90**
ACIC-2017 (21)	0.006	**0.004**	0.018	**0.017**	**1**	**1**
ACIC-2017 (22)	**0.004**	0.035	0.11	**0.096**	0.56	**0.81**
ACIC-2017 (23)	0.014	**0.01**	0.08	0.08	**0.97**	0.96
ACIC-2017 (24)	0.3	**0.04**	0.34	**0.25**	0.32	**0.90**
ACIC-2018 (1000)	290	**260**	4300	**3900**	0.57	**0.65**
ACIC-2018 (10000)	17	**13**	690	**680**	0.55	**0.68**
ACIC-2018 (2500)	**85**	99	1300	1300	0.59	**0.65**
ACIC-2018 (5000)	520	**510**	4100	4100	0.54	**0.64**
IHDP	**0.13**	**0.13**	0.46	0.46	**0.57**	**0.57**
Lalonde CPS	0.14	**0.084**	**0.22**	0.34	0.16	**0.75**
Lalonde PSID	**0.04**	**0.04**	**0.19**	0.44	0.46	**0.84**
Twins	0.22	**0.21**	0.24	**0.23**	0.51	**0.54**

Chapter 3

COMBINING T-LEARNER AND DR-LEARNING: A FRAMEWORK FOR ORACLE-EFFICIENT ESTIMATION OF CAUSAL CONTRASTS

Lars van der Laan, Marco Carone, and Alex Luedtke

Abstract

We introduce efficient plug-in (EP) learning, a novel framework for the estimation of heterogeneous causal contrasts, such as the conditional average treatment effect and conditional relative risk. The EP-learning framework enjoys the same oracle efficiency as Neyman-orthogonal learning strategies, such as DR-learning and R-learning, while addressing some of their primary drawbacks: (i) their practical applicability can be hindered by non-convex loss functions; and (ii) they may suffer from poor performance and instability due to inverse probability weighting and pseudo-outcomes that violate bounds. To overcome these issues, the EP-learner leverages an efficient plug-in estimator of the population risk function for the causal contrast. In doing so, it inherits the stability of plug-in strategies such as T-learning, while improving on their efficiency. Under reasonable conditions, EP-learners based on empirical risk minimization are oracle-efficient, exhibiting asymptotic equivalence to the minimizer of an oracle-efficient one-step debiased estimator of the population risk function. In simulation experiments, we show that EP-learners of the conditional average treatment effect and conditional relative risk outperform state-of-the-art competitors, including the T-learner, R-learner, and DR-learner. Open-source implementations of the proposed methods are available in our `R` package `hte3`.

3.1 Introduction

3.1.1 Background

In recent years, there has been a surge in interest in the development of tools designed for the flexible estimation of causally interpretable functions through the use of machine learning techniques. Many of these functions constitute subgroup-specific contrasts possibly describing heterogeneous treatment effects. Notable examples of such functions include the

conditional average treatment effect (CATE) [Robins, 2004, Kennedy, 2023a], conditional relative risk (CRR) [van der Laan et al., 2007, Richardson et al., 2017, Qiu et al., 2019], and causal conditional log-odds ratio [Vansteelandt and Goetghebeur, 2003, Robins and Rotnitzky, 2004, Tchetgen Tchetgen et al., 2010]. Under causal conditions, these functions can be expressed in terms of quantities estimable from randomly sampled data; often, these quantities involve outcome regressions, that is, conditional means of the outcome given certain baseline covariates and treatment levels.

Plug-in estimation, also known as T-learning [Künzel et al., 2019], offers a straightforward approach for estimating heterogeneous causal contrasts. In this approach, outcome regressions are estimated and then directly substituted into the contrast to compute the causal quantity of interest. However, a limitation of T-learners is that their performance is heavily dependent on the quality of the outcome regression estimators. This sensitivity is undesirable as the form of the outcome regressions can be significantly more complex—and therefore more difficult to estimate—than that of the causal contrast of interest. For example, the outcome regressions can be highly non-smooth, even if the conditional average treatment effect is constant. Consequently, there is a need for developing estimation methodologies that leverage the parsimony of the contrast function when feasible, while maintaining sufficient flexibility to accommodate complex functional forms [Luedtke and Laan, 2016, Kennedy, 2023a, Nie and Wager, 2021].

As a means of estimating a causal contrast, it is natural to first identify a risk function for this particular contrast. When an estimator for such risk function is available, contrasts can be estimated using flexible supervised learning tools such as regularized empirical risk minimizers [van de Geer, 2000], random forests [Breiman, 2001], and gradient boosting [Freund and Schapire, 1997], among others. However, estimating a risk function can be challenging, particularly when the loss function depends on unknown nuisance functions. For instance, the DR-learner loss [van der Laan, 2013, Luedtke and Laan, 2016, Kennedy, 2023a] and R-learner loss [Nie and Wager, 2021, Wager and Athey, 2018] for CATE estimation as well as the E-learner loss [Jiang et al., 2019] for CRR estimation depend on the outcome regression and propensity score. When a nonparametric efficient estimator of the resulting risk is used, the empirical risk minimizer based on estimated nuisance functions often estimates the causal contrast as accurately as the (oracle) empirical risk minimizer using the true nuisance functions [Wager and Athey, 2018, Foster and Syrgkanis, 2023, Kennedy, 2023a, Van Der Laan et al., 2023, Yang et al., 2023]. In parametric settings, an empirical risk minimizer based on an efficient risk estimator is typically efficient for the causal contrast under certain regularity conditions [van der Vaart, 1991, McClean et al., 2022, van der Laan et al., 2023]—this motivates the use of an efficient risk estimator in such problems. In nonparametric

settings, to our knowledge, such a general statement is not currently available. However, the advantages of using an efficient risk estimator have been theoretically established in the case of certain causal functions (Theorem 3 of Kennedy et al., 2017; Theorems 1 and 5 of Athey and Wager, 2021) and validated numerically for estimating the CATE [Kennedy, 2023a, Van Der Laan et al., 2023].

Many efficient estimators of commonly used causal contrast risks are based on loss functions that are Neyman-orthogonal [Robins et al., 1995, van der Laan and Robins, 2003, Chernozhukov et al., 2018a, Foster and Syrgkanis, 2023, Curth et al., 2020, Yang et al., 2023], including those leveraged by DR-learner and R-learner [Foster and Syrgkanis, 2023]. In most instances in the literature, a Neyman-orthogonal loss is derived, either explicitly or implicitly, by debiasing an initial plug-in risk estimator using the one-step estimation methodology [Pfanzagl and Wefelmeyer, 1985, Bickel et al., 1993]. However, there are two substantial drawbacks associated with efficient risk estimators built upon Neyman-orthogonal loss functions. First, Neyman-orthogonal loss functions can be nonconvex, even when they are derived by orthogonalizing a convex loss function. For instance, the Neyman-orthogonal losses associated with the E-learning risk for the CRR [Jiang et al., 2019, Qiu et al., 2019] and the bound-enforcing risk for the CATE [Luedtke and Laan, 2016] are nonconvex. Specifically, they can be expressed as weighted logistic regression loss functions with weights that may take negative values. This is especially problematic as numerous supervised learning implementations, including widely used software for fitting generalized additive models [Wood, 2011], random forests [Wright and Ziegler, 2017], and gradient-boosted regression trees [Chen and Guestrin, 2016], necessitate nonnegative weights. Second, in certain scenarios, empirical risk minimizers based on Neyman-orthogonal loss functions may yield extreme and unrealistic values for the causal contrast of interest. For example, when the outcome is binary, DR-learners [van der Laan, 2013, Kennedy, 2023a] may output treatment effect estimates residing outside of $[-1, 1]$, which is undesirable since the true CATE is confined within this interval. This failure to respect known bounds arises because fitting DR-learner involves regressing a pseudo-outcome on the covariates of interest, and this pseudo-outcome can take extreme values since it involves inverse-weighting by the estimated propensity score.

The inherent limitations of efficient risk estimators based on Neyman-orthogonal losses have been recognized in the literature, and several strategies have been proposed to address them. As these efficient risk estimators can yield nonconvex loss functions, researchers may opt for risk estimators derived from convex loss functions, even despite their inefficiency. Common examples include inverse-probability-weighted [Luedtke and Laan, 2016, Jiang et al., 2019, Chen et al., 2017] or plug-in [Curth and van der Schaar, 2021] risk estimators, which often fail to achieve parametric-rate consistency when the nuisance functions are

estimated flexibly. To avoid instability from inverse propensity weighting, overlap weighting can be used to down-weight observations in the tails of the propensity-score distribution [Li et al., 2019, Vansteelandt and Dukes, 2022, Morzywolek et al., 2023]. For example, IV-DR-learner uses a DR-learner loss with estimated overlap weights [Fisher, 2024], and R-learner uses an orthogonal loss for the overlap weighted risk (Corollary 9.1 of Robins, 2004; Section 5.2 of Robins et al., 2008; Nie and Wager, 2021). The key to the improved stability of these methods is that they more heavily weight regions of stronger overlap and underweight others. When the CATE is sufficiently smooth to allow reliable extrapolation from regions of greater overlap, such weighting can improve unweighted risk performance [Kennedy, 2023a, Kennedy et al., 2024, Fisher, 2024].

3.1.2 Our contributions

We provide a method for estimating causal contrasts based on a novel efficient plug-in risk estimator. Rather than focusing on a single risk function for a particular contrast, such as the CATE, we study a broad class of risk functions that includes this well-studied case as well as others, such as the CRR function. Like T-learners, EP-learners rely on the plug-in principle. However, unlike T-learners, they do this by computing a constrained minimizer of an efficient plug-in risk estimator. As such, EP-learners achieve the fast rates and oracle efficiency enjoyed by Neyman-orthogonal learning strategies. Our main contributions are as follows:

(i) we introduce the EP-learner for estimating causal contrasts, which is based on a novel efficient plug-in (EP) risk estimator;

(ii) we establish that, under reasonable conditions, the EP-learner is oracle-efficient, being asymptotically equivalent to the minimizer of an oracle-efficient one-step risk estimator;

(iii) we introduce EP-learners for the CATE and CRR functions, both of which are doubly robust.

3.2 Problem setup

3.2.1 Data structure and notation

Suppose that we have at our disposal a sample of n independent and identically distributed observations, $O_1, O_2, \ldots, O_n$, of the data structure $O = (W, A, Y)$ drawn from a probability distribution P_0 belonging to a nonparametric statistical model $\mathcal{M}$. In this data structure, $W \in \mathcal{W} \subset \mathbb{R}^d$ is a vector of baseline covariates, $A \in \mathcal{A} \subset \mathbb{R}$ is a possibly continuous treatment

assignment, and $Y \in \mathbb{R}$ is a bounded real-valued outcome. These observations may arise from an observational study or a randomized controlled trial. We let $\{Y^a : a \in \mathcal{A}\}$ be the set of potential outcomes associated with the observed data structure (W, A, Y) [Rubin, 2005], where Y^a is the outcome that would have been observed if, possibly contrary to fact, treatment $a \in \mathcal{A}$ had been administered.

For a given distribution $P \in \mathcal{M}$ and a generic realization (a, w) of (A, W), we denote the outcome regression by $\mu_P(a, w) := E_P(Y \mid A = a, W = w)$ and the propensity score by $\pi_P(a \mid w) := P(A = a \mid W = w)$. Further, we denote $\mu^a(w) := \mathbb{E}(Y^a \mid W = w)$, where $\mathbb{E}$ denotes the expectation under the joint distribution of the covariates W and the potential outcomes. Throughout, we let E_0^n denote the expectation with respect to the product measure P_0^n from which $(O_1, O_2, \ldots, O_n)$ is drawn. Further, we denote by $P_{0,W}$ the marginal distribution of W under P_0, and by $\|f\|$ and $\|f\|_\infty$ the $L^2(P_0)$ and P_0–essential supremum norm, respectively, of a given function $f \in L^2(P_0)$. For notational convenience, we denote the set $[m] := \{1, 2, \ldots, m\}$ for $m \in \mathbb{N}$ and write $\mathcal{S}_0$ to denote any summary $\mathcal{S}_{P_0}$ of the true distribution P_0.

3.2.2 Statistical goal

Our goal is to estimate a causal summary, such as the CATE, or its projection onto a convex action space $\mathcal{F}$. Specifically, we consider $\theta_0 = \operatorname{argmin}_{\theta \in \overline{\mathcal{F}}} R_0(\theta)$, the minimizer of a population risk function R_0 over the $L^2(P_{0,W})$–closure $\overline{\mathcal{F}}$ of $\mathcal{F}$. The function class $\mathcal{F}$ is user-specified and may be infinite-dimensional, as is the case for a reproducing kernel Hilbert space or Hölder class. If the causal summary lies in $\mathcal{F}$, then θ_0 coincides with the summary itself, with $\mathcal{F}$ encoding smoothness assumptions such as Hölder continuity. Otherwise, $\mathcal{F}$ serves as a working model that provides a useful approximation.

In this work, we restrict our attention to risk functions of the form $R_0 := R_{P_0}$, where $R_P(\theta) := E_P[L_{\mu_P}(W, \theta)]$ with loss function

$$L_{\mu_P}(w, \theta) := h_1(\theta(w)) \sum_{s \in \mathcal{A}} c_{s,1} g_1(\mu_P(s, w)) + h_2(\theta(w)) \sum_{s \in \mathcal{A}} c_{s,2} g_2(\mu_P(s, w)) \tag{3.1}$$

for arbitrary functions $g_1, g_2 : \mathbb{R} \to \mathbb{R}$ with a Lipschitz derivative, twice-differentiable Lipschitz functions $h_1, h_2 : \mathbb{R} \to \mathbb{R}$ with a continuous second derivative, and known constants $c_{a,1}, c_{a,2}$ for $a \in \mathcal{A}$. We introduce the general loss structure in (3.1) to encompass a wide range of causal summaries. Notably, any summary of the form $\theta_0 : w \mapsto \sum_{a \in \mathcal{A}} c_a f(\mu_a(w))$ with constants $c_a \in \mathbb{R}$ and fixed function $f : \mathbb{R} \to \mathbb{R}$ can be expressed in this manner. The particular form of the loss is not central to our method—beyond mild smoothness conditions, EP-learners can be constructed much more generally. In many applications, the loss simplifies considerably; for instance, in common cases g_1 and g_2 are simply the identity. The

CATE and CRR functions are notable examples of such summaries.

Example 3.1 (conditional average treatment effect)**.** The CATE $\theta_0^- : w \mapsto \mu_0(1,w) - \mu_0(0,w)$ minimizes the risk $\theta \mapsto R_0^-(\theta)$ over $L^2(P_{0,W})$, where we define

$$R_P^-(\theta) := E_P\left[\theta(W)^2 - 2\theta(W)\left\{\mu_P(1,W) - \mu_P(0,W)\right\}\right]. \tag{3.2}$$

This risk function is obtained by taking $\mathcal{A} = \{0,1\}$, $h_1(\theta) = \theta^2$, $h_2(\theta) = -2\theta$, $g_1 \circ \mu = 1$, $g_2 \circ \mu = \mu$, $c_{1,1} = 1$, $c_{0,1} = 0$, $c_{1,2} = 1$ and $c_{0,2} = -1$. We note that this same risk function can be used to learn any V-specific CATE function $v \mapsto E_0[\mu_0(1,W) - \mu_0(0,W) \,|\, V = v]$ for a coarsened covariate vector $V := f(W)$ with $f : \mathcal{W} \to \mathcal{V} \subseteq \mathbb{R}^{d_0}$ and $d_0 \leq d$. For example, V may represent a subset of components of W. This is achieved by minimizing $\theta \mapsto R_0^-(\theta)$ over $L^2(P_{0,V})$, where $P_{0,V}$ denotes the distribution of $f(W)$ under sampling from P_0, rather than $L^2(P_{0,W})$ [Morzywolek et al., 2023].

Example 3.2 (conditional relative risk)**.** When the outcome Y is binary or nonnegative, the log-CRR function $\theta_0^{\dot{-}} : w \mapsto \log \mu_0(1,w) - \log \mu_0(0,w)$ minimizes the risk $\theta \mapsto R_0^{\dot{-}}(\theta)$ over $L^2(P_{0,W})$ [Jiang et al., 2019, Qiu et al., 2019], where we define

$$R_P^{\dot{-}}(\theta) := E_P\left[\left\{\mu_P(1,W) + \mu_P(0,W)\right\}\log(1 + \exp\{\theta(W)\}) - \mu_P(1,W)\theta(W)\right]. \tag{3.3}$$

This risk function is obtained by taking $\mathcal{A} = \{0,1\}$, $h_1(\theta) = \log(1 + \exp(\theta))$, $h_2(\theta) = -\theta$, $g_1(\mu) = g_2(\mu) = \mu$, $c_{0,1} = c_{1,1} = 1$, $c_{0,2} = 0$ and $c_{1,2} = 1$. Similarly as in Example 3.1, for a coarsened covariate vector $V := f(W)$, this risk function can also be used to learn the V-specific CRR function $v \mapsto \log E_0[\mu_0(1,W) \,|\, V = v] - \log E_0[\mu_0(0,W) \,|\, V = v]$. A related exponential loss was also proposed by Chen et al. [2017].

We assume that the loss function giving rise to R_0 is γ–strongly convex [Equation 14.42 of Wainwright, 2019]—in Lemma B.15, we show that this condition suffices for the existence and uniqueness of θ_0. In Lemma B.16, we show that this strong convexity holds whenever there exists a constant $\gamma > 0$ such that $\sum_{a\in\mathcal{A}} c_{a,1} g_1(\mu_0(a,W)) \neq 0$, $\ddot{h}_2(\theta(W)) \neq 0$ and

$$\frac{\ddot{h}_1(\theta(W))}{\ddot{h}_2(\theta(W))} > -\frac{\sum_{a\in\mathcal{A}} c_{a,2} g_2(\mu_0(a,W))}{\sum_{a\in\mathcal{A}} c_{a,1} g_1(\mu_0(a,W))} + \frac{\gamma}{\ddot{h}_2(\theta(W)) \sum_{a\in\mathcal{A}} c_{a,1} g_1(\mu_0(a,W))}$$

both hold P_0–almost surely. Here, for $m \in \{1,2\}$, we denote the first derivatives of g_m and h_m as $\dot{g}_m$ and $\dot{h}_m$, respectively, and the second derivative of h_m by $\ddot{h}_m$. These conditions apply to both the CATE and CRR examples introduced above; see Examples 14.16 and 14.18 in Wainwright [2019] for details.

3.2.3 Our proposed approach: EP-learning

Given a population risk function R_0, it is natural to estimate the causal summary θ_0 using empirical risk minimization techniques based on an efficient estimator of R_0. In the existing literature, an 'orthogonal learning' strategy is often employed, wherein an efficient risk estimator is derived from a Neyman-orthogonal loss function [Foster and Syrgkanis, 2023]. Such a strategy benefits from relative insensitivity to the accuracy of involved nuisance estimators. In this section, we introduce, at a high level, our proposed EP-learning framework—an alternative approach to orthogonal learning—based on a novel, efficient plug-in estimator for the class of population risk functions under consideration.

To derive an efficient estimator of R_0, in the context of either orthogonal or EP-learning, we exploit the fact that the population risk parameter $P \mapsto R_P(\theta)$ at a specified $\theta \in \overline{\mathcal{F}}$ is a pathwise differentiable parameter under $\mathcal{M}$ and therefore amenable to nonparametric efficient estimation using standard techniques [Díaz and van der Laan, 2013]. Pathwise differentiability implies the existence of a nonparametric efficient influence function of the θ–specific population risk parameter, the variance of which provides the generalized Cramér-Rao lower bound for estimating $R_0(\theta)$. Specifically, under regularity conditions, for given $\theta \in \overline{\mathcal{F}}$, this efficient influence function has the form

$$D_{P,\theta} : (w, a, y) \mapsto L_{\mu_P}(\theta, w) + \Delta_{\pi_P,\mu_P}(w, a, y; \theta) - R_P(\theta) \ ,$$

where, letting $H_{m,\mu_P}(a, w) := c_{a,m} \cdot \dot{g}_m(\mu_P(a, w))$ for $m \in \{1, 2\}$, we write

$$\Delta_{\pi_P,\mu_P}(w, a, y; \theta) := \frac{1}{\pi_P(a \,|\, w)} \left\{ \sum_{m\in\{1,2\}} H_{m,\mu_P}(a, w) h_m(\theta(w)) \right\} \{y - \mu_P(a, w)\} \, .$$

In the theorem below, we formalize this fact. To do so, we require the following overlap condition:

A1) $P(\pi_P(a \,|\, W) > \delta) = 1$ for all $a \in \mathcal{A}$ and $P \in \mathcal{M}$.

Theorem 3.3. *Suppose Condition A1 holds. Then, for an arbitrary element $\theta \in \overline{\mathcal{F}}$, the nonparametric efficient influence function of $P' \mapsto R_{P'}(\theta)$ at $P \in \mathcal{M}$ is given by $D_{P,\theta}$.*

Knowledge of this efficient influence function is important because it encodes, in first order, the sensitivity of the population risk to perturbations in its nuisance parameters and facilitates the debiasing of plug-in estimators. We note that $D_{P,\theta}$ can be composed into a sum of three terms: the plug-in loss function L_{μ_P}, a weighted residual Δ_{π_P,μ_P} of the outcome regression μ_P, and the negative population risk $-R_P(\theta)$. Using that $E_P[D_{P,\theta}(O)] = 0$, this decomposition allows us to deduce that $L_{\mu_P} + \Delta_{\pi_P,\mu_P}$ is a Neyman-orthogonal loss function for R_P.

Let π_n and μ_n be estimators of π_0 and μ_0, respectively, which we assume for the time being are obtained using an independent dataset; later, we will describe approaches to

use cross-fitting when such an independent dataset is not available. Given these nuisance estimators, a one-step debiased estimator of the θ–specific risk $R_0(\theta)$ is given by

$$R_{n,\pi_n,\mu_n}(\theta) := \frac{1}{n}\sum_{i=1}^{n} L_{\mu_n}(\theta, W_i) + \frac{1}{n}\sum_{i=1}^{n} \Delta_{\pi_n,\mu_n}(O_i;\theta)\ . \tag{3.4}$$

This estimator can be seen as a debiased version of the plug-in estimator $\frac{1}{n}\sum_{i=1}^{n} L_{\mu_n}(\theta, W_i)$ with debiasing term $\frac{1}{n}\sum_{i=1}^{n} \Delta_{\pi_n,\mu_n}(O_i;\theta)$. Under appropriate conditions, $R_{n,\pi_n,\mu_n}(\theta)$ is an asymptotically efficient estimator of $R_0(\theta)$.

The decomposition in (3.4) suggests two approaches for learning θ_0 based on an efficient estimator of R_0. One approach, orthogonal learning, uses loss-based learning techniques based on the orthogonalized loss $L_{\mu_n} + \Delta_{\pi_n,\mu_n}$, that is, it relies on using the debiased risk estimator R_{n,π_n,μ_n}. The alternative approach we propose, EP-learning, instead relies on the *plug-in* risk estimator $R_n^* : \theta \mapsto \frac{1}{n}\sum_{i=1}^{n} L_{\mu_n^*}(\theta, W_i)$, where μ_n^* is a carefully constructed estimator of μ_0 that negates the need for the debiasing term $\frac{1}{n}\sum_{i=1}^{n} \Delta_{\pi_n,\mu_n^*}(O_i;\theta)$. When empirical risk minimization is used, the EP-learner for the causal contrast is the plug-in estimator $\operatorname{argmin}_{\theta\in\mathcal{F}} R_n^*(\theta)$ of the constrained risk minimizer $\theta_0 = \operatorname{argmin}_{\theta\in\mathcal{F}} R_0(\theta)$. In this sense, EP-learning follows the spirit of targeted minimum loss-based estimation (TMLE) [van der Laan and Rose, 2011, Luedtke et al., 2017].

Algorithm 3 Meta-algorithm for EP-learning (informal)

1: **Input:** Data $\{O_i\}_{i=1}^n$; nuisance estimators μ_n, π_n; function class $\mathcal{F}$; learning algorithm.

2: **Step 1: Debias nuisance.** Construct μ_n^* from μ_n and π_n so that the debiasing term

$$\tfrac{1}{n}\sum_{i=1}^{n} \Delta_{\pi_n,\mu_n^*}(O_i;\theta) = \tfrac{1}{n}\sum_{i=1}^{n} \big\{L_{\pi_n,\mu_n^*}(\theta, O_i) - L_{\mu_n^*}(\theta, W_i)\big\}$$

is small for all $\theta \in \mathcal{F}$.

3: **Step 2: Estimate risk by plug-in.** Define $R_n^*(\theta) := \frac{1}{n}\sum_{i=1}^{n} L_{\mu_n^*}(W_i;\theta)$ for $\theta \in \mathcal{F}$.

4: **Step 3: Estimate causal contrast.** Apply the learning algorithm to R_n^* to obtain an estimate θ_n^* of θ_0, e.g., $\theta_n^* := \operatorname{argmin}_{\theta\in\mathcal{F}} R_n^*(\theta)$.

5: **Output:** EP-learner θ_n^*

The EP-learner algorithm is designed to attain, on one hand, the oracle-efficiency of orthogonal learning strategies based on the loss $L_{\mu_n} + \Delta_{\pi_n,\mu_n}$, and on the other hand, the desirable properties—stability and loss convexity—enjoyed by plug-in estimation strategies based on the 'naive' loss L_{μ_n}. Rather than substituting any 'good' estimator of π_0 and μ_0 into the orthogonal loss function, in EP-learning, an outcome regression estimator μ_n^* is constructed from μ_n through a sieve-based adjustment to ensure that the debiasing term $\frac{1}{n}\sum_{i=1}^{n} \Delta_{\pi_n,\mu_n^*}(O_i;\theta)$ is negligible across values of θ, rendering explicit debiasing unnecessary.

Since the difference between R_n^* and R_{π_n,μ_n^*} is negligible, the EP-learner risk estimator benefits both from the plug-in property of the loss $L_{\mu_n^*}$ and the orthogonality property of the orthogonalized loss $L_{\mu_n^*}+\Delta_{\pi_n,\mu_n^*}$. We present a high-level meta-algorithm for EP-learning in Algorithm 3. For the class of losses we consider, the precise procedure for building the estimator μ_n^* is detailed later in Algorithm 4, and our EP-learner for θ_0 is presented in Section 3.4. In the next section, we illustrate how the plug-in property of EP-learning resolves the issues with Neyman-orthogonal learning referred to in the Introduction.

It is worth elaborating on the plug-in nature of EP-learning, particularly in contrast to T-learning. We say that an estimator $\widehat{R}_n(\theta)$ of the population risk $R_0(\theta)$ is a uniform plug-in estimator over $\theta \in \mathcal{F}$ if there exists a distribution $\widehat{P}_n \in \mathcal{M}$ such that $\widehat{R}_n(\theta) = R_{\widehat{P}_n}(\theta)$ for all $\theta \in \mathcal{F}$. The EP-learner risk estimator R_n^* is such a uniform plug-in estimator, with $\widehat{P}_n$ chosen so that its outcome regression satisfies $\mu_{\widehat{P}_n} = \mu_n^*$ and its marginal distribution of (W, A) equals the empirical distribution P_n. When paired with empirical risk minimization, EP-learner likewise returns a plug-in estimator, $\theta_n^* \in \operatorname{argmin}_{\theta\in\mathcal{F}} R_{\widehat{P}_n}(\theta)$, of the constrained minimizer $\theta_0 \in \operatorname{argmin}_{\theta\in\mathcal{F}} R_{P_0}(\theta)$. Similarly, T-learner returns a plug-in estimator of the unconstrained risk minimizer $\operatorname{argmin}_{\theta\in L^2(P_{0,W})} R_P(\theta)$, where the $\widehat{P}_n$ that is plugged in is any distribution whose outcome regressions equal μ_n; in the special case of CATE estimation, this unconstrained risk minimizer is just $\mu_n(1,\cdot) - \mu_n(0,\cdot)$ [Künzel et al., 2019]. Compared to the unconstrained risk minimizer, EP-learner exploits the restriction to $\mathcal{F}$ to inherit the favorable robustness properties of orthogonal learning. Indeed, when the initial estimator μ_n is misspecified, the EP-learner risk estimator can still be consistent uniformly over $\mathcal{F}$ provided the propensity is well-specified. As a consequence, EP-learner can still return a best approximation to the CATE in $\mathcal{F}$ in such cases.

While restricting to $\mathcal{F}$ may seem restrictive, this is standard in statistical learning theory [Vapnik, 1999, Foster and Syrgkanis, 2023, Wainwright, 2019]. In related problems such as regression, this choice is made to reflect structural or smoothness assumptions on the estimand. Such smoothness conditions are necessary because, without them, regression functions are not nonparametrically learnable [Vapnik and Alexey, 1971, Stone, 1982, Györfi et al., 2002]. For the CATE, the same limitation is reflected in minimax rates that become arbitrarily slow as smoothness decreases [Kennedy et al., 2024].

An efficient TMLE-based plug-in estimator for the population risk $R_0(\theta)$ at a specified candidate function θ for the causal dose–response curve was proposed in Díaz and van der Laan [2013]. However, their estimator does not provide a single plug-in estimator of the full population risk function R_0 and is therefore not amenable to empirical risk minimization over an infinite-dimensional function class. In contrast, the EP-learner risk estimator $R_n^*(\theta)$ provides a uniform plug-in estimator, meaning that the same outcome regression estimator

μ_n^* is used to compute $R_n^*(\theta)$ for every $\theta \in \mathcal{F}$. This uniform plug-in property ensures that if $\theta \mapsto R_P(\theta)$ is a convex risk function for any P, then $\theta \mapsto R_n^*(\theta)$ will also be convex. The concurrent work of Vansteelandt and Morzywołek [2023] similarly uses sieves to construct a debiased plug-in risk estimator, although they restrict their attention to estimation of a covariate-adjusted conditional mean. Their risk is a special case of (3.1) with $\mathcal{A} = \{1\}$, $h_1(\theta) = \theta^2$, $h_2(\theta) = -2\theta$, $g_1 \circ \mu = 1$, $g_2 \circ \mu = \mu$, $c_{1,1} = 1$, and $c_{0,1} = 0$. Besides considering a general class of risks, our approach uses a different form of cross-fitting and sieve-based adjustment to ensure the asymptotic equivalence of our EP-learner risk estimator with an oracle-efficient one-step risk estimator. This difference complicates our theoretical analysis, as it prohibits us from invoking statistical learning bounds based on sample-split nuisance estimators [Foster and Syrgkanis, 2023]. In particular, ensuring that the debiasing term $\frac{1}{n}\sum_{i=1}^n \Delta_{\pi_n,\mu_n}(O_i;\theta)$ is negligible (i.e., $o_p(n^{-1/2})$) requires the regression estimator μ_n^* to satisfy certain score equations on the full dataset (see Section 3.4 for details). This requirement is not unique to the EP-learner but is shared more broadly by sieve-based plug-in estimation [Shen, 1997, Chen, 2007, Qiu et al., 2021]. Consequently, the empirical process remainders in our analysis must be handled directly rather than through sample splitting, which necessitates sharper maximal inequalities and entropy bounds.

3.3 Limitations of existing Neyman-orthogonal learning strategies

3.3.1 Sensitivity of CATE DR-learner to large propensity weights

For estimation of the CATE, the DR-learner is a popular orthogonal learning approach based on the least-squares empirical risk function

$$\theta \mapsto \frac{1}{n}\sum_{i=1}^n \left\{\theta(W_i)^2 - 2\theta(W_i)\chi_n(W_i, A_i, Y_i)\right\},$$

where $\chi_n(w,a,y) := \mu_n(1,w) - \mu_n(0,w) + \frac{2a-1}{\pi_n(a|w)}\{y - \mu_n(a,w)\}$ is an estimated pseudo-outcome. This empirical risk coincides with the debiased risk estimator R_{n,π_n,μ_n} of the population risk R_0^- introduced in Example 3.1. A DR-learner can be obtained either by minimizing this empirical risk over $\mathcal{F}$ or by using standard regression methods to regress the pseudo-outcomes $\{\chi_n(W_i, A_i, Y_i)\}_{i=1}^n$ on the covariates $\{W_i\}_{i=1}^n$. As discussed in the Introduction, the pseudo-outcome used by the DR-learner can take extreme values when the propensity score is close to zero or one, which can lead to poor behavior of the resulting CATE estimator.

In contrast to the DR-learner, our EP-learner for the CATE—formally defined in Section 3.4—is based on the squared-error loss with the estimated pseudo-outcome $\mu_n^*(1,W) - \mu_n^*(0,W)$, where μ_n^* is an outcome regression estimator carefully constructed from the initial

estimates μ_n and π_n. Unlike a typical T-learner, the estimate μ_n^* is constructed to ensure that the resulting plug-in risk estimator

$$R_n^{*-}(\theta) := \frac{1}{n}\sum_{i=1}^{n}\left[\theta(W_i)^2 - 2\theta(W_i)\left\{\mu_n^*(1,W_i) - \mu_n^*(0,W_i)\right\}\right]$$

is efficient under reasonable conditions. EP-learning then estimates the CATE by a second-stage regression of $\{\mu_n^*(1,W_i) - \mu_n^*(0,W_i)\}_{i=1}^n$ on $\{W_i\}_{i=1}^n$. The combination of the specially constructed μ_n^* with the second-stage regression is what allows EP-learning to inherit the robustness properties of orthogonal learning strategies—such as fast rates and oracle efficiency—that the T-learner $\mu_n^*(1,\cdot) - \mu_n^*(0,\cdot)$ alone typically does not have. The construction of μ_n^* incorporates an inverse-propensity-weighted regression adjustment of μ_n, which renders the corresponding risk estimator doubly robust in the sense that, for each θ, $R_n^{*-}(\theta)$ is consistent if either the outcome regression or the propensity score is estimated consistently. This double robustness of the risk generally carries over to the corresponding EP-learner, as we show formally for empirical risk minimization in Section 3.5.

We now provide a heuristic explanation of why EP-learner can outperform DR-learner, particularly when paired with a local regression estimator, such as a Nadaraya-Watson [Nadaraya, 1964, Watson, 1964], K-nearest neighbors [Fix and Hodges, 1989], or random forest [Breiman et al., 1984] estimator. When used in a DR-learner or EP-learner, local methods estimate a regression function at a point by averaging the pseudo-outcomes of 'nearby' observations. When the number of points averaged is small, the estimates returned by DR-learner can be erratic due to the potentially large values taken by its estimated pseudo-outcomes. In contrast, the EP-learner is expected to be more stable since it is simply a local average of an initial CATE estimator. In fact, if this initial estimator is uniformly consistent, then the corresponding EP-learner remains consistent even when the number of local observations averaged is held fixed with sample size—a formal argument in the case of K-nearest neighbors estimation is provided in Section 3.6. Moreover, even if the initial estimator fails to be uniformly consistent, then EP-learner can still benefit from the desirable properties that it shares with DR-learner, namely the efficiency and double robustness of its risk estimator—details are provided in Section 3.5.1.

The benefits of EP-learner over DR-learner can be illustrated through a simple numerical experiment. To do so, we simulated a dataset consisting of 1,500 observations and used the `ranger` implementation of random forests [Wright and Ziegler, 2017] to estimate the CATE function based on both the DR-learner and EP-learner risks—the code used to run our experiment is provided in Appendix 3.B.1. In our illustration, the outcome is binary, and we use the known CATE range, $[-1, 1]$, to truncate DR-learner predictions. Figure 3.1a displays the estimated CATE curve for DR-learner and EP-learner with maximum tree depth

in random forests selected within $\{1, 2, \ldots, 7\}$ via 10-fold cross-validation. This figure reveals that the CATE curve estimated using EP-learner more accurately reflects the bell shape of the true CATE curve than using the DR-learner. This improvement is evident quantitatively, with a six-fold reduction in mean squared prediction error when using EP-learner (0.002) compared to DR-learner (0.012). To delve deeper into this phenomenon, we assessed the fits obtained by using four of the seven maximum tree depth values considered, namely 1, 2, 4 and 7 (Figure 3.1b). DR-learner fits become unstable at depths as low as 2, stemming from the limited observation numbers in the regression tree nodes and some pseudo-outcomes reaching magnitudes of 6—this can be seen in Figure 3.6 of Appendix 3.B.1. As a result, cross-validation chooses a depth of 1 for DR-learner, which fails to accurately capture the true CATE shape. In contrast, EP-learner fits tend to improve as the maximum depth increases from 1 to 7, with a selected maximum tree depth of 7. Regression trees, as used in random forests, are outcome-adaptive, often learned using CART [Breiman et al., 1984]. Thus, the difference in pseudo-outcomes between EP-learner and DR-learner affects not only the outcomes averaged within each tree node but also the learned tree structure itself. In particular, EP-learner uses the T-learner $\mu_n^*(1,\cdot) - \mu_n^*(0,\cdot)$ to build the tree structure, whereas DR-learner relies on a noisier pseudo-outcome. Thus, EP-learner can yield superior fits by leveraging less noisy pseudo-outcomes to both stabilize node averages and guide more accurate tree splits.

3.3.2 Nonconvexity of Neyman-orthogonal CRR loss function

We introduced in Example 3.1 the population risk function $R_0^{\dot{\div}}$ for the CRR $\theta_0^{\dot{\div}} : w \mapsto \log \mu_0(1, w) - \log \mu_0(0, w)$. In light of (3.4), a doubly-robust and efficient one-step debiased risk estimator for the population risk function is given by

$$R_{n,DR}^{\dot{\div}}(\theta) := \frac{1}{n} \sum_{i=1}^{n} \left[(\widehat{\mu}_{0,i} + \widehat{\mu}_{1,i}) \log \left(1 + \exp\{\theta(W_i)\}\right) - \widehat{\mu}_{1,i}\theta(W_i) \right], \tag{3.5}$$

where, for $s \in \{0, 1\}$ and $i \in [n]$, we define $\widehat{\mu}_{s,i} := \mu_n(s, W_i) + \frac{1(A_i=s)}{\pi_n(s \mid W_i)} \{Y_i - \mu_n(s, W_i)\}$. The corresponding doubly-robust orthogonal learner of the log-CRR function is computed by performing a weighted logistic regression of the pseudo-outcomes $\{\widehat{\mu}_{1,i}/(\widehat{\mu}_{0,i} + \widehat{\mu}_{1,i}) : i \in [n]\}$ on the covariate values $\{W_i : i \in [n]\}$ with weights $\{\widehat{\mu}_{0,i} + \widehat{\mu}_{1,i} : i \in [n]\}$.

Because the weights may take negative values, the Neyman-orthogonal loss associated with the one-step debiased risk estimator need not be convex. To illustrate, consider a simple example with a binary covariate $X \in \{0, 1\}$, an unrestricted parameter space $\Theta = \mathbb{R}^2$, and two observations ($n = 2$), one for each value of X. In this setting, the orthogonal loss takes the form

$$L_{\pi_n,\mu_n}^{\dot{\div}}(\theta_i, O_i) = \left(\widehat{\mu}_{0,i} + \widehat{\mu}_{1,i}\right) \Big[\log\{1 + \exp(\theta_i)\} - \tfrac{\widehat{\mu}_{1,i}}{\widehat{\mu}_{0,i} + \widehat{\mu}_{1,i}} \theta_i \Big],$$

so the second derivative with respect to θ_i is

$$\nabla^2 L^{\div}_{\pi_n,\mu_n}(\theta_i, O_i) = \big(\widehat{\mu}_{0,i} + \widehat{\mu}_{1,i}\big)\,\sigma(\theta_i)\big(1-\sigma(\theta_i)\big), \qquad \sigma(t) = \tfrac{e^t}{1+e^t}.$$

Since the weight $\widehat{\mu}_{0,i} + \widehat{\mu}_{1,i}$ may be negative, the Hessian can also be negative, implying that the loss is not guaranteed to be convex—even in this simple case. The same issue persists in higher-dimensional settings, where negative weights can lead to indefinite Hessians. Furthermore, the pseudo-outcomes may take values outside of $[0,1]$, which renders this a nonstandard use case for logistic regression. Figure 3.2 lists several common logistic regression implementations in `R` and indicates whether they allow the use of negative weights or outcomes that fall outside of $[0,1]$. As is revealed in this figure, most do not, which demonstrates the limited applicability of orthogonal learning in this setting. Figure 3.2 shows the first five rows of a synthetic input dataset containing the covariate values, pseudo-weights and pseudo-outcomes that would be entered into logistic regression software to estimate the log conditional relative risk. Notably, the mock dataset contains negative pseudo-weight values and pseudo-outcomes that fall outside of $[0,1]$. These issues are also encountered in the orthogonalization of the IPW loss for the CRR proposed in Section 2.3 of Chen et al. [2017]. Code to reproduce the mock dataset can be found in Appendix 3.B.1.

For a post-hoc constructed outcome regression estimator μ_n^*, our EP-learner of the CRR is based on the following EP risk estimator

$$R^{\div}_{n,EP}(\theta) := \frac{1}{n}\sum_{i=1}^{n}\{\mu_n^*(1,W_i)+\mu_n^*(0,W_i)\}\left[\log(1+\exp\{\theta(W_i)\}) - \frac{\mu_n^*(1,W_i)}{\mu_n^*(1,W_i)+\mu_n^*(0,W_i)}\theta(W_i)\right].$$

An estimator of the log-CRR function based on the EP-learner risk estimator is computed by performing the weighted logistic regression of the pseudo-outcome $\mu_n^*(1,W)/[\mu_n^*(1,W)+\mu_n^*(0,W)]$ onto the covariate vector W with weight $\mu_n^*(1,W)+\mu_n^*(0,W)$. Unlike for the efficient one-step risk estimator, the pseudo-weight is always nonnegative and the pseudo-outcome always falls in $[0,1]$. As a consequence, the EP-learner risk estimator corresponds with a convex loss function, and EP-learners of the log-CRR function can be implemented using most standard logistic regression software. In Section 3.4, we will provide an algorithm to construct μ_n^* so that $R^{\div}_{n,EP}$ is an efficient, doubly robust plug-in risk estimator.

3.4 Proposed approach: EP-learning algorithm

The EP-learner algorithm is designed to attain both the efficiency of the one-step debiased risk estimator in (3.4) and the desirable properties, such as stability and convexity, that are enjoyed by plug-in risk estimators. This learner is based on a risk function of the form $\theta \mapsto \frac{1}{n}\sum_{i=1}^{n} L_{\mu_n}(\theta, W_i)$ for an outcome regression estimator μ_n. Naturally, the fact that

EP-learner does not explicitly appear to debias this plug-in estimator may lead to concerns that it might be based on a suboptimal—in particular, inefficient—estimator of the risk. However, the EP-learner outcome regression estimator is carefully constructed to perform implicit debiasing. Specifically, it is designed such that the seemingly-critical debiasing term $\frac{1}{n}\sum_{i=1}^{n}\Delta_{\pi_n,\mu_n}(O_i;\theta)$ is negligible in an appropriate sense across values of θ, so that the plug-in risk estimator is asymptotically equivalent to a one-step debiased estimator. Before describing the general EP-learner approach, we illustrate its construction in the context of a simple, finite-dimensional example.

Example 3.1 (continued)**.** Consider the problem of CATE estimation from Example 3.1 in the special case where the action space $\mathcal{F}$ is the linear class including each function $w \mapsto \alpha^\top w$ with $\alpha \in \mathbb{R}^d$. In this finite-dimensional case, we write $R_0^-(\alpha)$ as shorthand for $R_0^-(w \mapsto \alpha^\top w)$. Given initial estimators π_n of π_0 and μ_n of μ_0, EP-learner uses the refined estimator $\mu_n^* : (a, w) \mapsto \mu_n(a, w) + (2a-1)\beta_n^\top w$, where

$$\beta_n := \operatorname*{argmin}_{\beta\in\mathbb{R}^d} \frac{1}{n}\sum_{i=1}^{n} \frac{1}{\pi_n(A_i \mid W_i)} \left\{Y_i - \mu_n(A_i, W_i) - (2A_i-1)\beta^\top W_i\right\}^2.$$

The first-order conditions satisfied by β_n ensure that the debiasing term for estimating $R_0^-(\alpha)$, notably $\frac{1}{n}\sum_{i=1}^{n}\frac{2A_i-1}{\pi_n(A_i\mid W_i)}\alpha^\top W_i\{Y_i - \mu_n^*(A_i, W_i)\}$, is exactly zero for $\alpha \in \mathbb{R}^d$. In this special case, the parametric EP-learner $w \mapsto (\alpha_n^*)^\top w$ based on μ_n^*, where $\alpha_n^* := \operatorname{argmin}_{\alpha\in\mathbb{R}^d}\sum_{i=1}^{n}\{\mu_n^*(1, W_i) - \mu_n^*(0, W_i) - \alpha^\top W_i\}^2$, corresponds to a targeted minimum loss-based estimator [van der Laan and Rose, 2011] of the best linear predictor of the CATE, $w \mapsto \alpha_0^\top w$, with $\alpha_0 := \operatorname{argmin}_{\alpha\in\mathbb{R}^d} E_0\{\mu_0(1, W) - \mu_0(0, W) - \alpha^\top W\}^2$.

When the dimension of the action space $\mathcal{F}$ is large relative to sample size (e.g., infinite action space), it is generally not possible to construct an outcome regression estimator μ_n that both has good predictive performance and makes the debiasing term $\frac{1}{n}\sum_{i=1}^{n}\Delta_{\pi_n,\mu_n}(O_i;\theta)$ exactly zero for each $\theta \in \mathcal{F}$. Indeed, there is an inherent trade-off between the size of the debiasing term and the mean squared error of the outcome regression estimator. The EP-learner algorithm is designed to carefully balance these two terms. It does so by separating the estimation of the outcome regression into two stages. In the first stage, statistical learning tools are used to obtain an initial estimate of the outcome regression with good predictive performance. In the second stage, this initial estimator is refined to make the debiasing term as small as possible without harming the performance of EP-learner.

The debiasing approach described in Example 3.1 can be generalized to infinite-dimensional function spaces using the idea of sieves. Specifically, if the linear span of a finite but growing set of basis functions approximates $\mathcal{F}$ increasingly well, then the refinement

procedure described in Example 3.1 could be performed within this span. This procedure would ensure that the debiasing term is small for each element in the linear space spanned by these basis functions. Critically, the number of basis functions used would need to grow with sample size so that the worst-case approximation error tends to zero. The separation of the outcome regression estimation and debiasing steps in the EP-learner ensures that the sieve does not need to approximate the true outcome regression μ_0 itself, but only needs to provide sufficient flexibility to reduce the debiasing term. For the CATE, given an approximating sequence of linear spaces $\mathcal{H}_k$ for $\mathcal{F}$ with dimension $k = k(n)$ typically chosen to grow with n, we can construct μ_n^* as in Example 3.1 so that the debiasing term $\frac{1}{n}\sum_{i=1}^n \Delta_{\pi_n,\mu_n^*}(O_i;\theta)$ is zero for each $\theta \in \mathcal{H}_k$. Assuming that $\mathcal{H}_k$ approximates $\mathcal{F}$ with vanishing error, we formally show in Section 3.5 that the empirical risk minimizer $\operatorname{argmin}_{\theta\in\mathcal{F}} \sum_{i=1}^n \{\mu_n^*(1,W_i) - \mu_n^*(0,W_i) - \theta(W_i)\}^2$ over $\mathcal{F}$ will inherit this debiasing property for the constrained minimizer $\operatorname{argmin}_{\theta\in\mathcal{H}_k} R_0(\theta)$.

We now formalize our general EP-learner algorithm, which is an instance of the meta-algorithm in Algorithm 3. Similar to competing procedures such as the DR-learner and R-learner, EP-learner can incorporate cross-fitting of the initial outcome regression and propensity score estimators to improve performance. Below, we present a general cross-fitted implementation of EP-learner. Let $J \in \mathbb{N}$ denote a fixed number of cross-fitting splits. Let $\mathcal{D}_n^1, \mathcal{D}_n^2, \ldots, \mathcal{D}_n^J$ be a partition of the available data $\mathcal{D}_n$ into J datasets of approximately equal size, corresponding to index sets $\mathcal{I}_n^1, \mathcal{I}_n^2, \ldots, \mathcal{I}_n^J$. For each $i \in [n]$, let $j(i) \in [J]$ be the index of the data fold containing observation i, so that $i \in \mathcal{I}_n^{j(i)}$. Suppose that, for each $j \in [J]$, we have constructed estimators $\pi_{n,j}$ of π_0 and $\mu_{n,j}$ of μ_0 based on $\mathcal{D}_n \backslash \mathcal{D}_n^j$. Let $\varphi_k : \mathbb{R}^d \mapsto \mathbb{R}^k$ be a feature mapping such that elements of the transformed action space $\mathcal{H} := \{h_j \circ \theta : \theta \in \mathcal{F}, j = 1, 2\}$ are approximated well by the linear space $\mathcal{H}_k := \{w \mapsto \beta^\top \varphi_k(w) : \beta \in \mathbb{R}^k\}$, where $k = k(n)$ is typically chosen to grow with n. Algorithm 4 is a particular choice of Step 1 of Algorithm 3. For each j, it transforms the initial outcome regression estimator $\mu_{n,j}$ into a refined estimator $\mu_{n,j}^*$ such that

$$\frac{1}{n}\sum_{i=1}^n \frac{1}{\pi_{n,j(i)}(A_i \mid W_i)} \left\{ \sum_{m\in\{1,2\}} H_{m,n,j(i)}(A_i, W_i)\psi_m(W_i) \right\} \{Y_i - \mu_{n,j(i)}^*(A_i, W_i)\} = 0 \tag{3.6}$$

for each $\psi_1, \psi_2 \in \mathcal{H}_{k(n)}$, where $H_{m,n,j}$ is defined as $(a, w) \mapsto c_{a,m}\dot{g}_m(\mu_{n,j}(a, w))$ with $c_{a,m}$ and g_m introduced in (3.1). We note that the EP-learner still relies on inverse propensity weight estimates through this step and, like the DR-learner, may suffer from instability when treatment overlap is limited.

Modifications of Algorithm 4 are possible. For certain population risk functions, including the CATE risk function used in Example 3.1, Algorithm 4 can be simplified by performing the adjustment over a lower-dimensional space; details are provided in Appendix 3.C.1. It is also possible to implement an alternative debiasing step that preserves bounds on the initial outcome regression estimator, even in cases in which the outcome itself may be unbounded;

details are provided in Appendix 3.C.2. Such a modification may improve performance by ensuring that the refinement step does not exceed bounds possibly specified a priori and enforced in the initial learning step.

Algorithm 4 Debiasing procedure for implementing Step 1 of Algorithm 3

Require:
- dataset $\mathcal{D}_n$ consisting of observations $(W_1, A_1, Y_1), (W_2, A_2, Y_2), \ldots, (W_n, A_n, Y_n)$;
- cross-fitted estimates $\pi_{n,1}, \pi_{n,2}, \ldots, \pi_{n,J}$ of π_0 and $\mu_{n,1}, \mu_{n,2}, \ldots, \mu_{n,J}$ of μ_0;
- feature mapping $\varphi_k : \mathbb{R}^d \mapsto \mathbb{R}^k$ of output dimension $k \in \mathbb{N}$;

1: for each $i \in [n]$, construct $\widehat{\varphi}_{k,j(i)} : \mathbb{R}^{d+1} \mapsto \mathbb{R}^{2k}$ as $(a, w) \mapsto (H_{1,n,i}(a, w), H_{2,n,i}(a, w))\varphi_k(w)$ with $H_{1,n,i}(a, w) := c_{a,1}\dot{g}_1(\mu_{n,j(i)}(a, w))$ and $H_{2,n,i}(a, w) := c_{a,2}\dot{g}_2(\mu_{n,j(i)}(a, w))$;

2: choose link function g and obtain regression coefficient estimate β_n as follows:

- *method 1: outcomes in* $[0, 1]$. set $g : x \mapsto \text{logit}(x)$ and $g^{-1} : x \mapsto \text{expit}(x)$, and compute coefficient vector estimate β_n obtained from logistic regression of outcome Y_i on feature vector $\widehat{\varphi}_{k,j(i)}(A_i, W_i)$ with offset $\text{logit}\, \mu_{n,j(i)}(A_i, W_i)$ and weight $1/\pi_{n,j(i)}(A_i \,|\, W_i)$ for $i \in [n]$;

- *method 2: general outcomes.* set $g : x \mapsto x$ and $g^{-1} : x \mapsto x$, and compute coefficient vector estimate β_n obtained from linear regression of outcome Y_i on feature vector $\widehat{\varphi}_{k,j(i)}(A_i, W_i)$ with offset $\mu_{n,j(i)}(A_i, W_i)$ and weight $1/\pi_{n,j(i)}(A_i \,|\, W_i)$ for $i \in [n]$;

3: for each $j \in [J]$, return debiased out-of-fold outcome regression estimator

$$\mu^*_{n,j} : (a, w) \mapsto g^{-1}\left(g(\mu_{n,j}(a, w)) + \widehat{\varphi}_{k,j}(a, w)^\top \beta_n\right). \tag{3.7}$$

We now present our EP-learner algorithm for feature mappings of fixed dimension k in Algorithm 5. This algorithm is a cross-fitted variant of Algorithm 3 that incorporates Algorithm 4 in Step 1. Line 2 of this algorithm encompasses many estimation strategies, including penalized empirical risk minimization [van de Geer, 2000], random forests [Breiman, 2001], and gradient boosting [Freund and Schapire, 1997]. Algorithm 6 instead describes a cross-validated EP-learner implementation incorporating data-driven selection of the feature mapping dimension. This selection is made based on a one-step debiased estimate of the risk function. The potential nonconvexity of this risk does not cause computational issues in this case since optimization is performed over a finite set of candidate values of k.

As presented in Algorithm 6, the cross-validated EP-learner only requires the construction of cross-fitted outcome regression and propensity score estimators once. Nuisance estimators do not need to be recomputed within each training fold. Since the cross-fitted

Algorithm 5 EP-learner algorithm (fixed sieve dimension)

Require:
- dataset $\mathcal{D}_n$ consisting of observations $(W_1, A_1, Y_1), (W_2, A_2, Y_2), \ldots, (W_n, A_n, Y_n)$;
- cross-fitted estimates $\pi_{n,1}, \pi_{n,2}, \ldots, \pi_{n,J}$ of π_0 and $\mu_{n,1}, \mu_{n,2}, \ldots, \mu_{n,J}$ of μ_0;
- feature mapping $\varphi_k : \mathbb{R}^d \mapsto \mathbb{R}^k$ of output dimension $k \in \mathbb{N}$;
- supervised learning algorithm for causal contrast;

1: obtain debiased outcome regression estimates $\mu^*_{n,j}$ from $\pi_{n,j}$ and $\mu_{n,j}$ using Algorithm 4;

2: return minimizer $\theta^*_{n,k}$ of $\theta \mapsto \frac{1}{n}\sum_{i=1}^n L_{\mu^*_{n,j(i)}}(\theta, W_i)$.

nuisance estimators are constructed using the entire dataset, there is some data leakage across the training folds used for cross-validation. To avoid such data leakage, the nuisance functions could alternatively be re-estimated within each training fold, although this would come at a possibly significant computational cost.

The sieve $\mathcal{H}_1, \mathcal{H}_2, \ldots$ used in the EP-learner implementation is an important ingredient that must be specified. In some cases, the transformed action space $\mathcal{H}$ suggests a natural choice of sieve. For instance, if $\mathcal{H}$ is a reproducing kernel Hilbert space, the linear span of the first k elements of the leading eigenfunctions of the integral kernel operator [Mendelson and Neeman, 2010] is a canonical choice for $\mathcal{H}_k$. More generally, sieves based on the trigonometric polynomials [Jackson, 1912], B-splines [Gordon and Riesenfeld, 1974], and wavelets [Antoniadis, 2007] are natural candidates with strong approximation guarantees under smoothness assumptions on the action space. In our experiments, we used the tensor-product trigonometric polynomial sieve proposed by Zhang and Simon [2023] and implemented in the associated R package `Sieve` [Zhang, 2023]. In principle, the cross-validated EP-learner could be modified to select from different basis functions used to define the sieve. For example, Algorithm 5 can be used to construct EP-learners based on different subsets of trigonometric polynomial and B-spline basis functions.

3.5 Theoretical guarantees for empirical risk minimization

3.5.1 Efficiency of the EP risk estimator

In this section, we study the EP risk estimator presented in Algorithm 5 with a deterministic feature mapping dimension $k = k(n)$ growing with sample size. Specifically, we establish conditions under which this EP risk estimator is equivalent to an oracle-efficient one-step risk estimator. To do so, we show that the debiasing term for the EP-learner risk estimator

Algorithm 6 EP-learner algorithm (cross-validated sieve dimension)

Require:
- dataset $\mathcal{D}_n$ consisting of observations $(W_1, A_1, Y_1), (W_2, A_2, Y_2), \ldots, (W_n, A_n, Y_n)$;
- number of cross-fitting splits J;
- growing sequence $\varphi_1, \varphi_2, \ldots, \varphi_{K(n)}$ of feature mappings, with $\varphi_k : \mathbb{R}^d \mapsto \mathbb{R}^k$ and $K(n) \in \mathbb{N}$;
- supervised learning algorithm for causal contrast;

partition $\mathcal{D}_n$ into J datasets $\mathcal{D}_n^1, \mathcal{D}_n^2, \ldots, \mathcal{D}_n^J$ of approximately equal size, say with corresponding index sets $\mathcal{I}_n^1, \mathcal{I}_n^2, \ldots, \mathcal{I}_n^J$;
for $j \in [J]$, construct estimates $\pi_{n,j}$ and $\mu_{n,j}$ of π_0 and μ_0 using $\mathcal{D}_n \setminus \mathcal{D}_n^j$;
for $k = 1, 2, \ldots, K(n)$ **do**
  for $j = 1, 2, \ldots, J$ **do**
    obtain debiased estimates $\mu^*_{n,j,k}$ from $\pi_{n,j}$ and $\mu_{n,j}$ using Algorithm 4 with dataset $\mathcal{D}_n \backslash \mathcal{D}_n^j$, cross-fitted estimates $\pi_{n,1}, \pi_{n,2}, \ldots, \pi_{n,J}$ and $\mu_{n,1}, \mu_{n,2}, \ldots, \mu_{n,J}$, and feature-mapping φ_k;
    obtain minimizer $\theta^*_{n,j,k}$ of $\theta \mapsto \sum_{i \in [n] \backslash \mathcal{I}_n^j} L_{\mu^*_{n,j(i)}}(\theta, W_i)$.
  end for
end for
compute optimal dimension $k_{cv}(n) := \operatorname{argmin}_{k \in \mathbb{N}} \frac{1}{n} \sum_{i=1}^n L_{\pi_{n,j(i)}, \mu_{n,j(i)}}(O_i, \theta^*_{n,j(i),k})$;
return $\theta^*_{n,k_{cv}(n)}$ obtained using Algorithm 5 with dataset $\mathcal{D}_n$ and feature mapping $\varphi_{k_{cv}(n)}$.

has a second-order dependence on the estimation error of the outcome regression estimator and the sieve approximation error of the transformed action space $\mathcal{H}$. Thus, if this debiasing term converges to zero rapidly enough, the EP-learner risk estimator is $n^{\frac{1}{2}}$–consistent, asymptotically linear, and nonparametric efficient for the population risk uniformly over the action space $\mathcal{F}$, as we formalize below.

In what follows, we say 'the loss is linear in μ' if both g_1 and g_2 are the identity map in (3.1), and is not linear in μ otherwise. For $\theta \in L^2(P_{0,W})$, we denote by $\|\theta\|_\infty$ the P_0–essential supremum norm and by $\Pi_k \theta \in \operatorname{argmin}_{\phi \in \mathcal{H}_k} \|\theta - \phi\|$ the $L^2(P_0)$–projection of $\theta \in \mathcal{H}$ onto the sieve $\mathcal{H}_k$. Below, we make use of the following conditions:

C1) there exist constants $\eta \in (0, 1)$ and $M \in (0, \infty)$ such that, for all $j \in [J]$, for P_0–almost every realization (w, a) of (W, A), and with P_0–probability tending to one:

(a) *strong positivity:* $\pi_0(a \,|\, w) \in [\eta, 1 - \eta]$ and $\pi_{n,j}(a \,|\, w) \in [\eta, 1 - \eta]$;

(b) *boundedness:* $|\mu_{n,j}(a, w)| < M$ and $|\mu^*_{n,j}(a, w)| < M$;

C2) *polynomial nuisance rates:* there exist $\gamma > 0$ and $\beta > \frac{1}{4\gamma}$ such that, for all $j \in [J]$:

(a) *propensity score:* $\|\pi_{n,j} - \pi_0\| = \mathcal{O}_p\big(n^{-\gamma/(2\gamma+1)}\big)$;

(b) *outcome regression:* $\|\mu_{n,j} - \mu_0\| = \mathcal{O}_p\big(n^{-\beta/(2\beta+1)}\big)$. Moreover, if the loss is not linear in μ, then $\beta > 1/2$, so that $\|\mu_{n,j} - \mu_0\| = o_p(n^{-1/4})$.

(c) *debiased outcome regression:* $\|\mu^*_{n,j} - \mu_0\| = \mathcal{O}_p\big(n^{-\beta/(2\beta+1)} + \sqrt{k(n)\log n/n}\big)$;

C3) *sufficiently small sieve approximation error:* there exist $\rho > \frac{1}{2}$ and $\nu \geq 0$ such that:

(a) *slowly growing Lebesgue constant:* $\sup_{\phi\in\mathcal{H}, v\in\mathcal{H}_{k(n)}} \frac{\|\Pi_{k(n)}(\phi+v)\|_\infty}{\|\phi+v\|_\infty} = \mathcal{O}\big(\{\log k(n)\}^\nu\big)$;

(b) *polynomial approximation rate:* $\sup_{\theta\in\mathcal{H}} \inf_{\phi\in\mathcal{H}_{k(n)}} \|\theta - \phi\|_\infty = \mathcal{O}\big(k(n)^{-\rho}\big)$;

(c) *large enough exponent:* if $\min\{\beta,\gamma\} > \frac{1}{2}$, then $\rho \geq \frac{1}{2\beta+1}$, and otherwise, $\rho \geq \frac{1}{4\min\{\beta,\gamma\}}$.

Conditions C1 and C2 of Theorem 3.2 commonly appear in the nonparametric inference literature [Bickel et al., 1993, Shen, 1997, van der Laan and Robins, 2003, van der Laan and Rose, 2011, Chernozhukov et al., 2018a]. Condition C1a strengthens A1 so that strong positivity also holds for the propensity score estimators. Conditions C2a and C2b are satisfied for some $\gamma > 1/2$ and $\beta > 1/2$ if the cross-fitted nuisance estimators converge at a polynomial rate faster than $n^{-1/4}$. Under suitable smoothness assumptions, such rates can be achieved by several statistical learning methods, including generalized additive models [Hastie and Tibshirani, 1987], reproducing kernel Hilbert space estimators [Nie and Wager, 2021], the highly adaptive lasso [van der Laan, 2017], and certain neural network architectures [Farrell et al., 2018]. For the CATE and log-CRR risks in Section 3.3, the loss is linear in μ, Condition C2 requires that the rate exponent of the outcome regression estimators satisfy $\beta > 1/(4\gamma)$. Condition C2a and C2b together represent a rate double robustness condition [Bang and Robins, 2005, Rotnitzky et al., 2021]. Condition C2c imposes a high-level rate requirement on the sieve-adjusted estimator $\mu^*_{n,j}$. In particular, under C2b it holds if $\max_{j\in[J]} \|\mu^*_{n,j} - \mu_0\| = \mathcal{O}_p(\max_{j\in[J]} \|\mu_{n,j} - \mu_0\| + \sqrt{k(n)\log n/n})$, where the first and second terms on the right capture the sieve approximation error for μ_0 and the statistical estimation error of a $k(n)$-dimensional parameter, respectively [Chen, 2007]. In Appendix B.5, we verify this condition explicitly for the case where $(\mu^*_{n,j} : j \in [J])$ is obtained from Method 2 in Algorithm 4. Condition C4 controls the complexity of the action space $\mathcal{F}$ in supremum norm. This condition holds with exponent $\alpha := s/d$ when $\mathcal{F}$ falls in a d–variate Hölder or Sobolev smoothness class with smoothness exponent $s > d/2$ (Theorem 2.7.2. of van der Vaart and Wellner, 1996; Corollary 4 of Nickl and Pötscher, 2007).

By controlling the Lebesgue constant [Belloni et al., 2015] in C3a, approximation theory can be used to bound $\sup_{\theta\in\mathcal{H}} \|\theta - \Pi_{k(n)}\theta\|_\infty$ by the sup-norm approximation error of C3b

up to logarithmic dependence on $k(n)$ [Huang, 2003, Chen and Christensen, 2013, Belloni et al., 2015]. Only requiring elements of $\mathcal{H}$ to be continuous, Section 3.2 of Belloni et al. [2015] provides an overview of several sieve choices that satisfy this condition, including those based on trigonometric series, splines, wavelets, and local polynomials. Condition C2c is the convergence rate expected for an empirical risk minimizer over a sieve space of dimension $k(n)$ based on a loss indexed by cross-fitted nuisances [Foster and Syrgkanis, 2023]. This condition is satisfied under regularity conditions when each $\mu^*_{n,j}$ is obtained using Algorithm 4. Together, conditions C2 and C3 ensure the existence of a sequence of sieve dimensions $k(n)$ such that, for each $j \in [J]$, remainder terms of the order $\|\mu^*_{n,j} - \mu_0\| \, \|\pi_{n,j} - \pi_0\|$ and $\sup_{\phi \in \mathcal{H}} \|\phi - \Pi_{k(n)}\phi\| \|\mu^*_{n,j} - \mu_0\|$ are $\mathcal{O}_p(n^{-\frac{1}{2}})$. The precise choice of $k(n)$ depends on β, γ and ρ, which are typically unknown. However, under mild conditions, using arguments similar to those in van der Laan and Dudoit [2003], it can be shown that the cross-validated EP-learner is equivalent to the infeasible EP-learner that uses an oracle selector of $k(n)$ minimizing the population risk.

Compared to cross-fitted empirical risk estimators [Kennedy, 2023a, Foster and Syrgkanis, 2023], the nuisance estimator $\mu^*_{n,j}$ depends on the entire dataset, and so, the EP-learner loss function $L^*_{\mu_{n,j}}$ is not independent of observations in the jth fold. Consequently, to obtain tighter control of empirical process remainders [van de Geer, 2014], we require that the action space $\mathcal{F}$ not be too large with respect to the supremum norm metric. To quantify the complexity of $\mathcal{F}$, we consider the ε–covering number $N_\infty(\varepsilon, \mathcal{F})$ with respect to the $P_{0,W}$–essential supremum metric $(\theta_1, \theta_2) \mapsto d_\infty(\theta_1, \theta_2) := \|\theta_1 - \theta_2\|_\infty$ on $\mathcal{F}$ as the smallest number of balls of d_∞–radius ε required to cover $\mathcal{F}$ [Chapter 2 of van der Vaart and Wellner, 1996], and the corresponding metric entropy integral $\mathcal{J}_\infty(\delta, \mathcal{F}) := \int_0^\delta \{\log N_\infty(\varepsilon, \mathcal{F})\}^{\frac{1}{2}} d\varepsilon$. To establish our results, we require the following regularity conditions on the class of functions $\mathcal{F}$:

C4) *nonparametric, convex action space that is not too large:* $\mathcal{F}$ is uniformly bounded, convex, and $\mathcal{J}_\infty(\delta, \mathcal{F}) \leq C\delta^{1-1/(2\alpha)}$ for some $\alpha > 1/2$ and $C > 0$, and for every $\delta > 0$.

In the following theorem, $l^\infty(\mathcal{F})$ denotes the Banach space of bounded functionals $b : \mathcal{F} \to \mathbb{R}$ equipped with the norm $b \mapsto \|b\|_{\ell^\infty(\mathcal{F})} := \sup_{\theta \in \mathcal{F}} |b(\theta)|$. Finally, for $\theta \in \mathcal{F}$, we denote the EP-learner risk estimator obtained from Algorithm 5 by $R_{n,k(n)}(\theta) := \frac{1}{n} \sum_{i=1}^n L_{\mu^*_{n,j(i)}}(\theta, W_i)$.

Theorem 3.2 (Oracle efficiency of EP-learner risk)**.** *Suppose that conditions C1–C4 hold and that $k(n)$ is a deterministic sequence satisfying the rate conditions*

$$n^{\frac{1}{2\rho(2\beta+1)}} \cdot \frac{(\log n)^{\nu/\rho}}{k(n)} \longrightarrow 0, \qquad \frac{k(n) \log n}{n^{\frac{2c(\beta,\gamma)}{2c(\beta,\gamma)+1}}} \longrightarrow 0,$$

with $c(\beta,\gamma) := \min\{\beta,\gamma,1/2\}$. *Then, it holds that*

$$n^{\frac{1}{2}} \sup_{\theta\in\mathcal{F}} \left| \frac{1}{n}\sum_{i=1}^{n} \Delta_{\pi_{n,j(i)},\mu^*_{n,j(i)}}(O_i;\theta) \right| \xrightarrow{p} 0 \ .$$

Furthermore, it follows that $\sup_{\theta\in\mathcal{F}} |n^{\frac{1}{2}}\{R_{n,k(n)}(\theta) - R_{n,0}(\theta)\}| \xrightarrow{p} 0$, *and thus, the stochastic process* $(n^{\frac{1}{2}}\{R_{n,k(n)}(\theta) - R_0(\theta)\} : \theta \in \mathcal{F})$ *converges weakly in* $\ell^\infty(\mathcal{F})$ *to the tight Gaussian process* $\mathbb{G}$ *with mean zero and covariance function* $(\theta_1,\theta_2) \mapsto P_0(D_{0,\theta_1}D_{0,\theta_2})$.

We note that the constraints on γ, β and ρ imposed by C2 and C3 ensure the existence of a sieve dimension $k(n)$ satisfying the growth rate bounds of Theorem 3.2. Since $D_{0,\theta}$ is the (nonparametric) efficient influence function of $P \mapsto R_P(\theta)$, it holds that $R_{n,k(n)}(\theta)$ is a nonparametric efficient estimator of $R_0(\theta)$ for each $\theta \in \mathcal{F}$ [Bickel et al., 1993]. In addition to being efficient, the estimator $R_{n,k(n)}$ exhibits a form of double robustness whenever the loss is linear in μ; this is the case for both the CATE and log-CRR loss functions introduced in Sections 3.3.1 and 3.3.2. In particular, even when the outcome regression estimators are inconsistent, the EP-learner risk estimator $R_{n,k(n)}(\theta)$ is still consistent for $R_0(\theta)$ if $\|\pi_{n,j} - \pi_0\|$ and $\sup_{\phi\in\mathcal{H}} \|\phi - \Pi_{k(n)}\phi\|$ tend to zero in probability—this is studied in Lemmas B.12 and B.14 of the Supplement.

Theorem 3.2 indicates that no additional correction is needed to ensure the asymptotic linearity and efficiency of the plug-in risk function estimator whenever it is based on the debiased outcome regression estimators $\mu^*_{n,1}, \mu^*_{n,2}, \ldots, \mu^*_{n,J}$. However, this property comes at the cost of condition C3, which requires that the product of the outcome regression estimator rate $\|\mu^*_{n,j} - \mu_0\|$ and sieve approximation error $\sup_{\phi\in\mathcal{H}} \|\phi - \Pi_{k(n)}\phi\|_\infty$ be $\mathcal{O}_p(n^{-\frac{1}{2}})$. Fixing nuisance rate exponents β and γ of C2, in order to obtain a sieve approximation rate satisfying C3, elements in the transformed action space $\mathcal{H}$ must generally be sufficiently smooth with respect to the sieve basis. If $\mathcal{H}$ is a subset of the class of d–variate Hölder smooth functions with smoothness exponent $s > d/2$, this condition is known to hold with $\rho := s/d$ for appropriate sieves based on trigonometric, spline, wavelets, and local polynomial basis functions (Section 2.3 of Chen, 2007; Section 3 of Belloni et al., 2015). For appropriate sieves based on trigonometric and wavelet basis functions, condition C3 also holds with $\rho := s/d$ when $\mathcal{H}$ is contained in the d–variate Sobolev class of smoothness exponent $s > d/2$ [Kon and Raphael, 2002, Cobos et al., 2016].

3.5.2 Convergence rate guarantees for ERM-based EP-learners

We now establish properties of an EP-learner of the population risk minimizer $\theta_0 := \operatorname{argmin}_{\theta\in\mathcal{F}} R_0(\theta)$ constructed using empirical risk minimization. We refer to this special case of Algorithm 5 as the ERM-based EP-learner $\theta^*_{n,k(n)}$ defined as any element

of $\operatorname{argmin}_{\theta \in \mathcal{F}} R_{n,k(n)}(\theta)$. To simplify our results, we assume that $\theta^*_{n,k(n)}$ always exists. However, when that is not the case, it suffices to identify a near-minimizer in the sense of Foster and Syrgkanis [2023].

To theoretically evaluate the performance of the ERM-based EP-learner, we compare it with the oracle learner $\theta_{n,0} \in \operatorname{argmin}_{\theta \in \mathcal{F}} R_{n,\pi_0,\mu_0}(\theta)$, which minimizes the oracle efficient one-step risk estimator over $\mathcal{F}$. Before studying the ERM-based EP-learner, we present an upper bound on the rate of the oracle learner. For the following theorem and the remainder of this section, the exponent $\alpha > 1/2$ is the entropy integral exponent from C4.

Theorem 3.3 (Oracle learner rate)**.** *Under C1a and C4, it holds that* $\|\theta_{n,0} - \theta_0\| = \mathcal{O}_p(n^{-\alpha/(2\alpha+1)})$.

If $\mathcal{F}$ consists of d–variate, compactly-supported, Hölder-smooth functions with Hölder exponent $s > d/2$, the entropy integral satisfies C4 with $\alpha = s/d$ [Theorem 2.7.2. of van der Vaart and Wellner, 1996], which gives an oracle rate of $n^{-\alpha/(2\alpha+1)}$. When the target of interest is the value of the CATE function at a point, Kennedy et al. [2024] shows that this rate is minimax for an oracle learner based on observing counterfactual outcomes. The results in Yang and Barron [2000] can similarly be used to show that this oracle rate is minimax when the goal is instead estimation of the CATE function in an $L^2(P_{0,W})$–sense. Hence, at least in the CATE estimation setting, this suggests that the rate for Hölder smooth functions provided by Theorem 3.3 cannot be improved. In light of this, throughout, we will refer to $n^{-\alpha/(2\alpha+1)}$ as the oracle rate.

The following theorem provides an upper bound on the convergence rate of the ERM-based EP-learner. To obtain fast rates for the ERM-based EP-learner, we impose the following coupling between the supremum and $L^2(P_0)$ norms:

C5) *sup-norm coupling:* there exists $C > 0$ such that $\|\theta\|_\infty \leq C \|\theta\|^{1-1/(2\alpha)}$ for all $\theta \in \mathcal{F} - \mathcal{F}$.

Condition C5 is known to hold with exponent $\alpha := 1/d$ when $\mathcal{F}$ is a subset of the d–variate Hölder smoothness class of order $s = 1$ [Lemma 4 of Bibaut et al., 2021]. Moreover, this condition holds for appropriate reproducing kernel Hilbert spaces [Lemma 5.1 of Mendelson and Neeman, 2010]; in particular, it holds with exponent $\alpha := s/d$ when $\mathcal{F}$ is a subset of the d–variate Sobolev smoothness class of order $s > d/2$. This condition also holds when $\mathcal{F}$ equals the signed convex hull of appropriate basis functions [Lemma 2 of van de Geer, 2014]. For empirical risk minimization within an RKHS, such sup-norm couplings have been used to establish fast rates under L^2 convergence of the nuisance estimators [Nie and Wager, 2021, Foster and Syrgkanis, 2023].

Our next result describes how close the ERM-based EP-learner and its oracle counterpart are, and involves the sequence $\varepsilon_n^2 := \min\{a_n, b_n\}$ with

$$a_n := \frac{(\log n)^{2\nu}}{k(n)^{\rho}} \left[\frac{1}{n^{\beta/(2\beta+1)}} + \left\{ \frac{k(n) \log n}{n} \right\}^{1/2} + \left\{ \frac{k(n)^{(\rho/\alpha)}}{n} \right\}^{1/2} \right];$$
$$b_n := \frac{1}{n^{2\beta/(2\beta+1)}} + \frac{k(n) \log n}{n} + \frac{(\log n)^{2\nu}}{n^{1/2} k(n)^{\rho\{1-1/(2\alpha)\}}} .$$

Theorem 3.4 (EP-learner convergence rate for a deterministic sieve growth rate)**.** *Suppose that conditions C1, C2, C3a , C3b, C4 and C5 hold. Then, it holds that*

$$\|\theta^*_{n,k(n)} - \theta_{n,0}\| = \mathcal{O}_p\left(\varepsilon_n\right) + \mathcal{O}_p\left(n^{-\alpha/(2\alpha+1)}\right)$$

for any sieve growth rate satisfying $(\log n)k(n)/n^{2c(\beta,\gamma)/\{2c(\beta,\gamma)+1\}} \to 0$.

Theorem 3.4 establishes that, under certain conditions, the rate of the root-mean-squared error between the EP-learner and the oracle learner is dominated by the oracle rate up to a remainder term ε_n that depends on nuisance estimation rates and the sieve approximation error. In view of the b_n term, the EP-learner rate is no worse than the debiased outcome regression estimator rate $n^{-\beta/(2\beta+1)} + \{k(n) \log n/n\}^{1/2}$ as long as (i) θ_0 is at least as smooth as μ_0 so that $\beta \leq \alpha$; and (ii) the sieve growth rate is fast enough so that $n^{\beta/\{\rho(2\beta+1)\}}(\log n)^{2\nu\beta/\rho}/k(n) \to 0$. The rate in Theorem 3.4 does not appear to depend directly on γ or β, except through the sieve approximation error term ε_n. This is because C2 ensures that the errors of the initial nuisance estimators can be absorbed into the $\mathcal{O}_p(n^{-\alpha/(2\alpha+1)})$ term.

The next result demonstrates that, under appropriate conditions, the EP-learner is oracle-efficient in the sense of Kennedy [2023a] so long as the oracle rate is tight in expectation and the sieve approximation error and nuisance estimation rates tend to zero quickly enough. To achieve oracle efficiency, we require $k(n)$, β and ρ to be such that $n^{\alpha/(2\alpha+1)}\varepsilon_n \to 0$, which mandates that $k(n)$ grow at a faster rate than imposed by Theorem 3.2. We also require the following additional condition:

C6) *tight oracle learner rate:* $n^{-\alpha/(2\alpha+1)}/E_0^n \|\theta_{n,0} - \theta_0\| = \mathcal{O}(1)$.

Theorem 3.5 (Oracle efficiency of ERM-based EP-learner)**.** *Suppose that conditions C6, C2 with* $\gamma > 1/2$ *and* $\beta > \max\{\frac{1}{4\gamma}, \frac{2\alpha}{4\alpha^2+1}\}$*, and C3b with* $\rho = \alpha$ *hold in addition to the conditions of Theorem 3.4. Then,*

$$\|\theta^*_{n,k(n)} - \theta_{n,0}\| = \mathcal{O}_p\left(E_0^n \|\theta_{n,0} - \theta_0\|\right)$$

for any sieve growth rate satisfying $\frac{(\log n)\, k(n)}{n^{2c(\beta,\gamma)/(2c(\beta,\gamma)+1)}} \to 0$ *and* $a_n n^{\frac{2\alpha}{2\alpha+1}} \to 0$.

By the reverse triangle inequality, Theorem 3.5 implies that $\|\theta^*_{n,k(n)} - \theta_0\| - \|\theta_{n,0} - \theta_0\|$ tends to zero in probability faster than $E_0^n\|\theta_{n,0} - \theta_0\|$. If this statement can be strengthened to convergence in expectation rather than in probability, it would follow that $E_0^n\|\theta^*_{n,k(n)} - \theta_0\|/E_0^n\|\theta_{n,0} - \theta_0\| \to 1$. In such case, $\theta^*_{n,k(n)}$ would converge in expected mean squared error at the same rate and constant as the oracle learner.

In view of Lemma B.20 in Appendix B.4.4, there exists a sieve growth rate $k(n)$ satisfying the growth bounds of Theorem 3.5. It is not immediately obvious that it is possible to obtain a sieve approximation rate exponent ρ equal to the metric entropy and a supremum norm coupling exponent α satisfying C4 and C5, as is required in Theorem 3.4. However, this is indeed guaranteed if, for example, $\mathcal{F}$ and $\mathcal{H}$ consist of d–variate functions that are Hölder smooth with order $s = 1$ or Sobolev smooth with order $s > d/2$.

3.5.3 Discussion and related literature

Under the conditions of Theorem 3.5, EP-learner is oracle efficient if both nuisance estimators converge in L^2 at polynomial rates faster than $n^{-1/4}$ ($\gamma > 1/2$ and $\beta > 1/2$). These rate conditions are analogous to those for orthogonal learning when no additional structure is assumed beyond C4 and C5 (Theorem 3 of Nie and Wager, 2021; Example 1 of Foster and Syrgkanis, 2023). We expect that alternative conditions can be obtained by working with L^4 rates [cf. Foster and Syrgkanis, 2023, Proposition 1]. Here, we instead follow Nie and Wager [2021] in formulating conditions in terms of L^2 rates, as these are more widely established.

In doubly robust settings, DR and other orthogonal learners can achieve oracle efficiency when the outcome regression converges more slowly than $n^{-1/4}$, provided the propensity is estimated sufficiently quickly so that the product of the L^2 errors is $o_p(n^{-1/2})$. In this regime, EP-learner's outcome regression can similarly converge slower than $n^{-1/4}$, but the rate requirement is dictated by the more stringent of the usual product rate condition ($\beta > 1/[4\gamma]$) and a condition dictated by the smoothness of the causal summary ($\beta > 2\alpha/[4\alpha^2+1]$). The latter requirement is imposed to ensure the sieve approximation error is negligible. As $\alpha \to 1/2$, this condition is at its most restrictive, requiring an $o_p(n^{-1/4})$ rate of convergence for the outcome regression. However, once α is large enough (more than $2\gamma - 1/[8\gamma]$), the usual product rate condition is the more restrictive, and so the requirement on the outcome regression rate is the same as for orthogonal learners.

The rates achieved by many Neyman-orthogonal learning strategies—including the DR-learner, R-learner, and EP-learner—may be suboptimal when additional structure is imposed on the causal contrast and nuisance functions. In particular, for CATE estimation under Hölder smoothness of the estimand, the propensity score, and the outcome regression, these learners fail to attain the minimax rate and oracle efficiency in low-smoothness regimes of

the nuisances. By contrast, specialized procedures that exploit Hölder regularity and higher-order bias corrections can achieve minimax-optimal rates in such settings [Kennedy, 2023a, Kennedy et al., 2024, Fisher, 2024]. However, even for marginal treatment effect estimation, such rate improvements are known to not be achievable in structure-agnostic settings, where only high-level L^2 rate conditions are imposed on the nuisances [Balakrishnan et al., 2023, Jin and Syrgkanis, 2024].

3.6 Example of when EP-learners are more robust than DR-learners: K-nearest neighbors

In this section, we provide theoretical support for how EP-learners can outperform orthogonal learners, such as DR-learners, by being less sensitive to tuning parameter choices, as discussed in Section 3.3.1. Focusing on estimation of the CATE using K-nearest neighbors, we show that if the adjusted T-learner $\mu_n^*(1,\cdot) - \mu_n^*(0,\cdot)$, used to construct pseudo-outcomes for EP-learning, is uniformly consistent, then EP-learner remains consistent even with a fixed number of neighbors as n grows.

Consider estimation of the CATE based on the population risk function in Section 3.3.1. Let μ_n^* be constructed from initial estimators π_n and μ_n of π_0 and μ_0 using Algorithm 4, where for each $j \in [J]$ we set $\pi_{n,j} = \pi_n$ and $\mu_{n,j} = \mu_n$. Let $\theta_{n,EP}$ denote the K_n-nearest neighbors EP-learner, defined pointwise for each $w \in (0,1)^d$ as

$$\theta_{n,EP}(w) = \frac{1}{K_n} \sum_{i=1}^{n} 1\{i \in \mathcal{I}_{K_n,n}(w)\} \left[\mu_n^*(1, W_i) - \mu_n^*(0, W_i)\right],$$

where $\mathcal{I}_{K_n,n}(w)$ denotes the indices of the K_n nearest observations to w under the Euclidean norm, with ties broken arbitrarily. Our main result is the following theorem.

Theorem 3.6. *Suppose $W_1, \ldots, W_n$ are independent draws from the uniform distribution on $[0,1]^d$ and that, for each $a \in \{0,1\}$, $w \mapsto \mu_0(a,w)$ is Lipschitz-continuous on $[0,1]^d$. At each sample size n, let the number of neighbors K_n be an arbitrary element of $[n]$. Then,*

$$\sup_{w \in (0,1)^d} \left|\theta_{n,EP}(w) - \theta_0^-(w)\right| \lesssim \sup_{w \in (0,1)^d} \left|\mu_n^*(1,w) - \mu_n^*(0,w) - \theta_0^-(w)\right| + \mathcal{O}_p\left((K_n \log n/n)^{1/d}\right).$$

Importantly, the above establishes consistency provided $K_n = o(n/\log(n))$ and $\mu_n^*(1,\cdot) - \mu_n^*(0,\cdot)$ is uniformly consistent; a special case is that $K_n = 1$ for all n. Hence, for K-nearest neighbors, EP-learner is more robust to suboptimal tuning than the DR-learner, as illustrated in Figure 3.1, where it remains stable across a wider range of tuning parameters.

3.7 Numerical experiments

3.7.1 Evaluating EP-learner

We evaluated the empirical performance of the implementation of our cross-validated EP-learner described in Algorithm 6 through three experiments based on the CATE and log-CRR risk functions introduced in Sections 3.3.1 and 3.3.2.

In the first two experiments, we considered estimation of the CATE function. In the first experiment, the covariate vector W was taken to have dimension $d = 3$, whereas in the second, it was taken to have dimension $d = 20$ but with only 5 active components. The goal of the second experiment is to assess how EP-learner performs under sparsity of the causal summary. Because the sieve uses all covariates without penalization, it is natural to ask whether including basis functions for inactive variables degrades performance. In the third experiment, we considered estimation of a log-CRR with a covariate vector of dimension $d = 3$. In each experiment, we further considered four settings characterized by estimation based on a simple versus moderately complex functional form, and on limited versus moderate propensity overlap across treatment arms. Additional details on the design of these simulation experiments are provided in Appendix 3.B.2.

In all experiments conducted, the data-generating mechanism implied a causal contrast additive in the covariates, and the EP-learner sieve was constructed using a univariate trigonometric cosine polynomial basis, as considered in Zhang and Simon [2023]. The truncation level for the trigonometric series was selected from the first six leading frequencies of the trigonometric series using cross-validation. To obtain initial cross-fitted propensity score and outcome regression estimates, ensemble learning based on cross-validation and a library of candidate algorithms including gradient-boosted trees (xgboost) with tree depths ranging from 1 to 8 and generalized additive models (GAMs) was used. The performance of EP-learner was compared to that of DR-learner, R-learner [Nie and Wager, 2021], T-learner, and a causal forest estimator [Wager and Athey, 2018] for the CATE in the first two experiments. We considered various EP-learners, DR-learners, T-learners, and R-learners, each differing in the supervised learning algorithm used in Step 2 of Algorithm 5. The supervised learning algorithms considered were GAMs, multivariate adaptive regression splines (MARS), random forests (RF), and xgboost, and are implemented using the `R` packages `gam` [Hastie and Hastie, 2015], `earth` [Milborrow, 2019], `ranger` [Wright and Ziegler, 2017], and `xgboost` [Chen and Guestrin, 2016]. We note that only GAM (for a fixed tuning parameter) is an example of an empirical risk minimizer, with $\mathcal{F}$ taken as the cubic smoothing spline class [Hastie and Tibshirani, 1990], whereas all other methods are greedy estimators that select tree splits or basis functions in an adaptive manner. For the R-learner, we omitted the

use of MARS because the weight implementation of `earth` is computationally prohibitive. The causal forest estimator used is implemented using the R package `grf` and was internally tuned using the tune.parameters="all" argument of the `causal_forest()` function [Tibshirani et al., 2018]. For the DR-learner, R-learner, T-learner, and EP-learners, tuning parameters were selected using 10-fold cross-validation based on the DR-learner loss.

In the third experiment, we compare the performance of EP-learner for log-CRR estimation with that of the T-learner and of the IPW-based E-learner introduced in Jiang et al. [2019]. We did not consider learners based on the one-step debiased risk estimator due to the nonconvexity of its corresponding loss function. The base-learner MARS was omitted for all causal learners as the `R` package `earth` does not support non-binary outcomes for logistic regression. For similar reasons, we used the `xgboost` (rather than the `ranger`) implementation of random forests. Apart from these changes, the base statistical learning algorithms used for the causal learners were the same as those used for CATE estimation.

For each experimental setup and sample size $n \in \{500, 1000, 2000, 3000, 4000, 5000\}$, we generated 1000 simulated datasets. The Monte Carlo empirical mean squared error was computed for each EP-learner and its competitors. Results for the CATE experiments with a complex functional form and moderate treatment overlap are displayed in Figure 3.3, whereas results for the log-CRR experiment with limited and moderate treatment overlap are shown in Figure 3.4. Results for simple causal contrasts and settings with limited treatment overlap are qualitatively similar and summarized in Appendix 3.B.2.

In most experimental settings and simulation scenarios, EP-learner performed at least as well as—and in some cases better than—all other methods considered, particularly when using tree-based methods. In the case of CATE learners based on GAMs, EP-learner and DR-learner had comparable performance across all sample sizes. However, the R-learner and T-learner outperformed EP-learner and DR-learner for GAM with $d = 3$, suggesting that plug-in estimation and overlap weighting were beneficial for improving efficiency in this case. For more flexible algorithms, such as tree-based learners and MARS, EP-learner significantly outperformed DR-learner in mean squared error across all settings, even for large sample sizes. For instance, in the CATE experiments with `xgboost`, EP-learner reduced the mean squared error approximately three-fold compared to DR-learner at sample size $n = 5000$. In all CATE experiments, R-learner and causal forests performed well. In the CATE experiment with covariate dimension $d = 3$, the `ranger` and `xgboost` EP-learners exhibited better mean squared error than R-learner and causal forests across all settings and sample sizes. In the CATE experiment with covariate dimension $d = 20$, causal forests slightly outperformed the `ranger` and `xgboost` EP-learners for smaller sample sizes, but EP-learners perform best for larger sample sizes. In the log-CRR experiments, we found that EP-learners generally

performed as well or better than the T-learner, depending on the degree of treatment overlap. The IPW-learner exhibited significantly worse performance than EP-learner in most settings.

3.7.2 Comparing with R-learner

We reproduce the four benchmark settings (A–D) from Nie and Wager [2021], where covariates X are drawn from a distribution P_d, treatment is assigned via a propensity $e^\star(x)$, and outcomes follow $Y = b^\star(X) + (W - \frac{1}{2})\tau^\star(X) + \sigma\varepsilon$ with $W \mid X \sim \text{Bernoulli}(e^\star(X))$ and $\varepsilon \sim \mathcal{N}(0,1)$. Setup A features complex nuisances but a simple treatment effect; Setup B mimics a randomized trial with no confounding; Setup C imposes strong confounding with an easy propensity, a difficult baseline, and a constant treatment effect; and Setup D generates uncorrelated treated and control outcomes so joint learning offers no benefit. We fix $d = 5$ and $\sigma = 1$, vary $n \in \{500, 1000, 2000\}$, and use the exact data-generating functions from Nie and Wager [2021].

Figure 3.5 reports the mean squared error (MSE) of CATE estimators across sample sizes $n = 500, 1000, 2000$ under settings A–D, averaged over 100 independent data draws. The EP-learner consistently outperformed or matched the performance of the DR-learner and R-learner across all settings when the base learner was XGBoost or Random Forest, while all methods performed similarly with GAM. The EP-learner also outperformed the T-learner except in Setup D, where the strong performance of the T-learner is expected since the treated and control models are unrelated and separate fitting naturally aligns with the data-generating process. With random forests, EP-learner was competitive with causal forests (the random-forest–based R-learner from `grf`): it outperformed them in Setups B and D and matched their performance in larger samples for Setup A. Setup C features a constant CATE and overlap weighting used by the R-learner and causal forests is expected to be beneficial since extrapolation from regions of high overlap to those of lower overlap is straightforward. This is reflected in R-learner and causal forests generally outperforming both EP-learner and DR-learner. Overall, EP-learner either outperforms or matches the DR-learner and R-learner with gradient boosting or random forests across diverse regimes, while remaining competitive with state-of-the-art causal forests.

3.8 Conclusion

In this work, we have introduced a framework for estimating a heterogeneous causal contrast based on a particular construction of an efficient plug-in risk function estimator. To the best of our knowledge, our construction gives the first doubly-robust, efficient estimator of a population risk function for the log-CRR function that corresponds to a convex loss function. While our primary focus has been on EP-learners for use in settings with a single

time-point, discrete treatment, our framework can be extended to contexts involving continuous or longitudinal treatments. In future research, it would be of interest to explore the development of EP-learners for broader causal summaries, such as conditional treatment effects under longitudinal interventions [Luedtke et al., 2017, Rotnitzky et al., 2017]. In contrast to the plug-in methods presented in these earlier works, EP-learner would make it possible to leverage parsimony found in contrasts between counterfactual means under longitudinal interventions but otherwise not found in the counterfactual means themselves, for example.

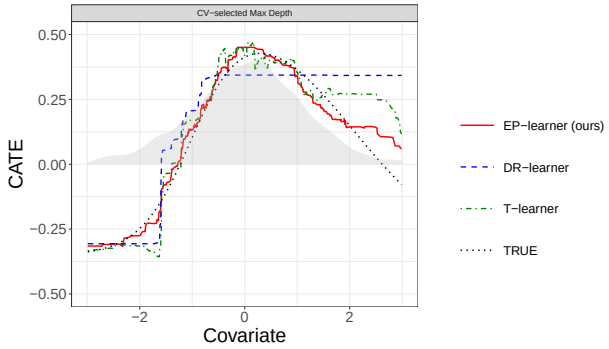
To determine sieve specifications for EP-learner, including basis function types and dimensions, we proposed using the cross-validated one-step risk estimator. In high-dimensional settings, defining a candidate basis for constructing a sieve can be challenging. In our simulation experiments, additive sieves based on univariate trigonometric series performed well, even with complex action spaces like random forests and boosted trees. However, incorporating tensor-product sieves with higher interaction degrees could potentially improve the performance of EP-learner. Alternatively, it may be fruitful to data-adaptively learn a subset of variables that drive treatment effect heterogeneity and then use basis functions of those variables. In the case of estimating the covariate-adjusted conditional mean, Vansteelandt and Morzywołek [2023] proposes a penalized sieve method for constructing debiased plug-in risk estimators. Adapting our approach to incorporate penalization would be an interesting area for future work.

While in this work we assumed access to an estimator of the propensity, from which estimates of the inverse propensity were obtained, we note that EP-learner can also be modified to use inverse-propensity weights estimated directly—for example, using stable balancing weights [Zubizarreta, 2015] or autoDML [Chernozhukov et al., 2022c]. Moreover, EP-learner may benefit from stabilizing propensity score estimates through calibration techniques [van der Laan et al., 2024b, Klaassen et al., 2025].

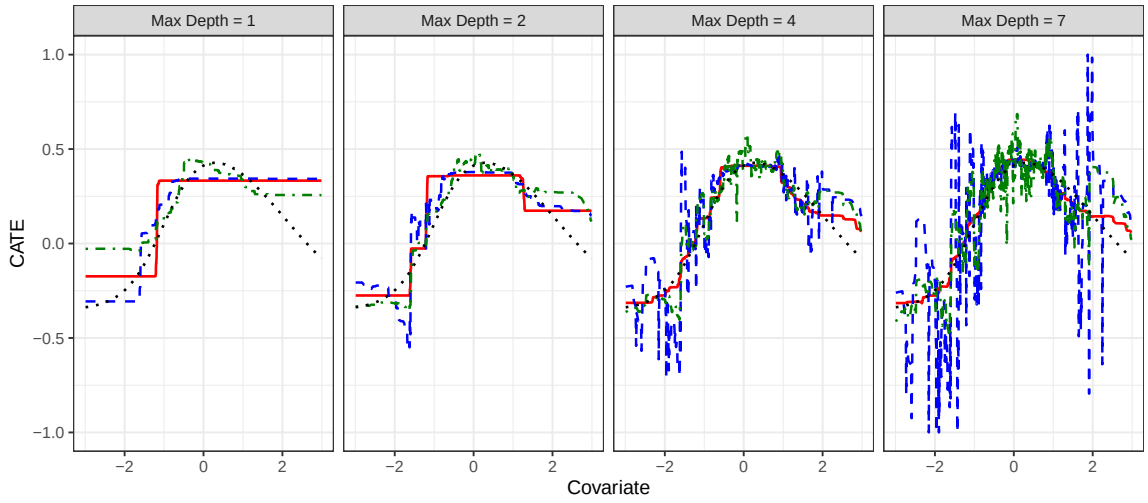
An unexplored application of EP-learner is its potential to improve the calibration of predictors for the CATE through causal isotonic calibration [Van Der Laan et al., 2023]. Isotonic calibration often exhibits poor calibration at the boundary of uncalibrated predictions, primarily because of overfitting of isotonic regression in that region [Groeneboom and Lopuhaa, 1993]. However, this issue could be mitigated by utilizing an EP-learner risk estimator for the CATE, which would ensure that the pseudo-outcome employed for calibration itself serves as a CATE estimator.

Acknowledgements. Research reported in this publication was supported by NIH grants DP2-LM013340 and R01-HL137808, and NSF grant DMS-2210216. The content is solely

the responsibility of the authors and does not necessarily represent the official views of the funding agencies.



(a) Cross-validated maximum tree depth



(b) Fixed maximum tree depth

Figure 3.1: CATE estimates based on EP-learner, DR-learner, and T-learner with random forests and various maximum tree depths computed on a single dataset. (Top) EP-learner, DR-Learner, and T-learner CATE estimates with 10-fold cross-validated maximum tree depth; mean squared prediction errors are 0.0021, 0.012, and 0.005, respectively. Observed covariate distribution is depicted in gray. (Bottom) EP-learner, DR-learner,and T-learner CATE estimate for maximum tree depths of 1, 2, 4 and 7.

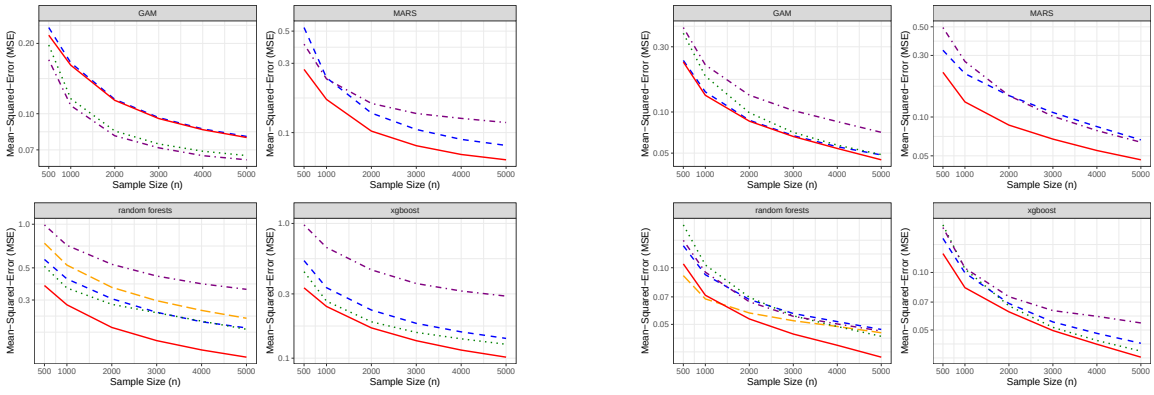
R Package	Accepts negative weights	Accepts outcomes outside of [0,1]
earth	✗	✗
gam/mgcv	✗	✗
glm	✗	✗
hal9001	✗	✗
ranger	✗	✗
randomForest	✗	✗
xgboost with default loss (∗)	✗	✗
xgboost with custom weighted loss (∗)	✓	✓

Negative pseudo-weights

Covariate	Pseudo-weight	Pseudo-outcome
1.0333	1.57	0.66
-1.755	**-1.94**	**-0.09**
0.637	2.17	0.41
1.980	**-4.04**	**-0.20**
-0.735	1.58	**1.15**

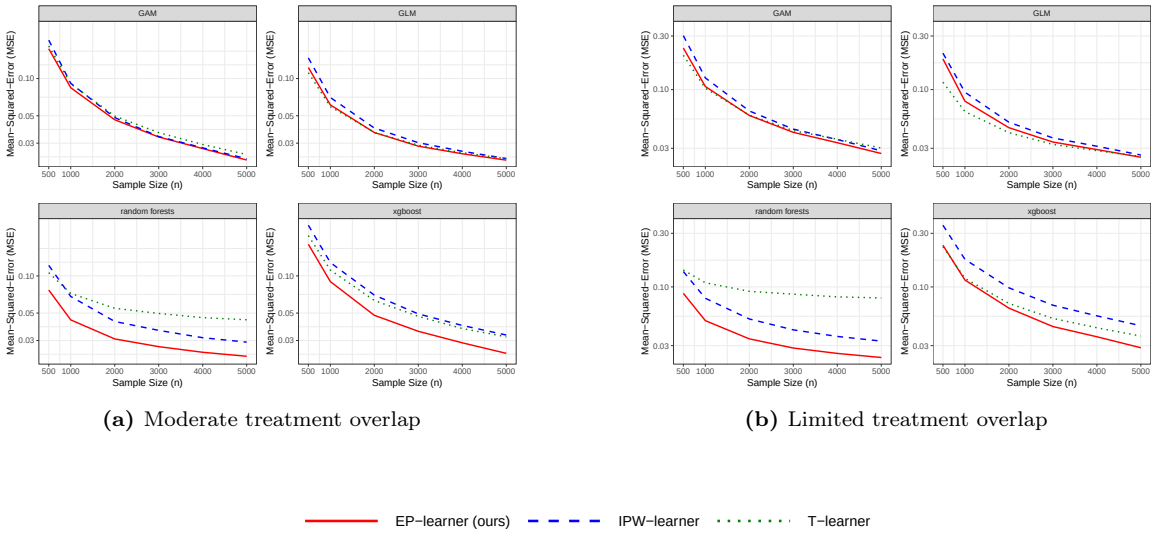
Pseudo-outcome falls outside [0,1]

Figure 3.2: (Left) Common R packages for logistic regression may or may not allow negative weights and outcomes outside of [0,1]; a checkmark indicates that they do. (∗) The xgboost package has many built-in loss functions, but they do not accept negative weights. However, these negative weights can be absorbed into a custom loss function. (Right) Example dataset for estimating the CRR using the one-step efficient risk estimator.



(a) $d = 3$ **(b)** $d = 20$ with 5 active

Figure 3.3: CATE experiments with a complex CATE and moderate treatment overlap: Mean squared error for DR-learner, R-learner, T-learner, and cross-validated EP-learner with supervised learning algorithm GAM, MARS, `ranger`, and `xgboost`. For the plots reporting the results of `ranger`, we also display the results of causal forests for comparison.



(a) Moderate treatment overlap

(b) Limited treatment overlap

Figure 3.4: CRR experiments with a complex CRR and moderate (left) and limited (right) treatment overlap: Mean-squared error for IPW-learner, T-learner, and CV-EP-learner with supervised learning algorithm GAM, random forests, and xgboost. The DR-learner algorithm for the CRR is not implemented due to nonconvexity of the loss.

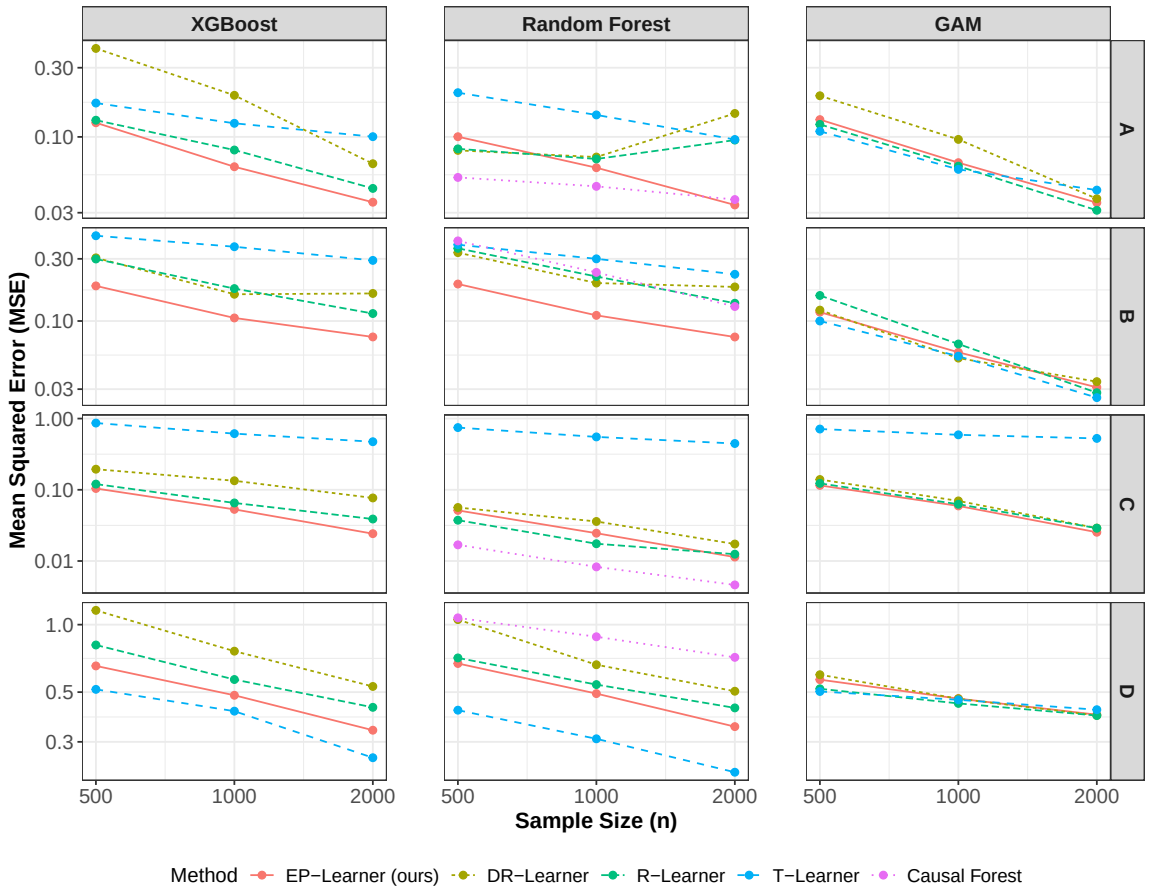


Figure 3.5: Mean squared error (MSE) of CATE estimates across sample sizes $n \in \{500, 1000, 2000\}$ under the four simulation settings (A–D) of Nie and Wager [2021], comparing EP-learner, R-learner, DR-learner, and T-learner across different base learners (XGBoost, Random Forest, GAM). For comparison, the causal forest implementation from `grf` is also displayed under the Random Forest base learner.

Chapter Appendix

3.A Code

An `R` package, *hte3*, implementing the EP-learner is available on GitHub at: https://github.com/Larsvanderlaan/hte3. Code to reproduce our experiments can be found here: https://github.com/Larsvanderlaan/hte3/tree/main/paper_EPlearner_experiments.

3.B Supplemental information for experiments

3.B.1 Small-scale experiments

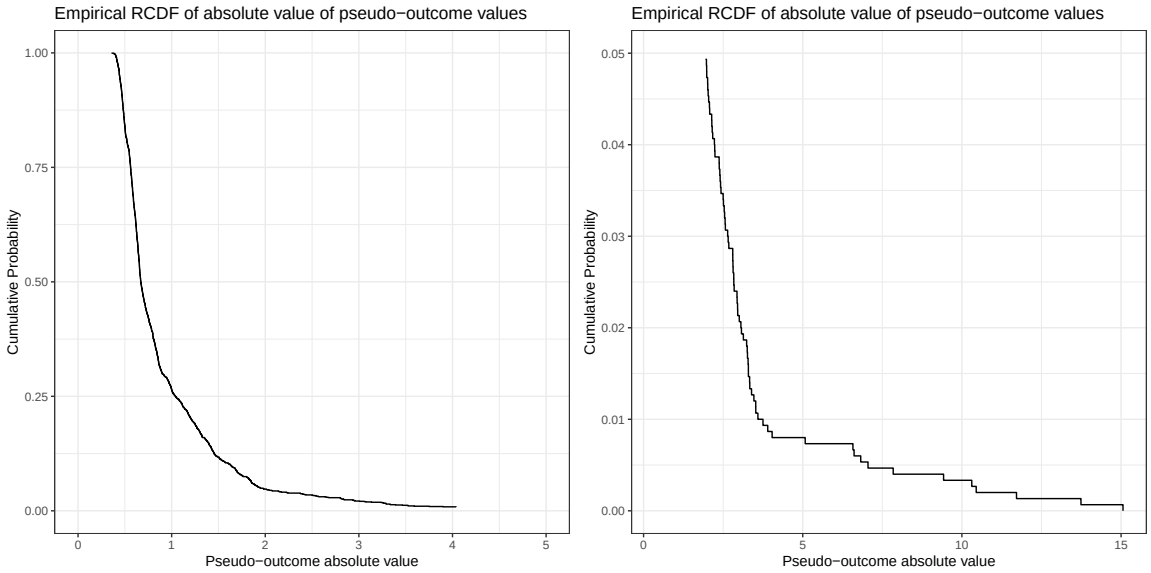


Figure 3.6: (Left) Empirical reverse cumulative distribution function (RCDF) of estimated DR pseudo-outcome values for the simulated dataset. (Right) A zoomed-in empirical RCDF. The 1% and 99% empirical quantiles of the pseudo-outcome are -3.2 and 3.1 while the minimum and maximum values are -16 and 14.

The following `R` code was used to generate the dataset corresponding to Figure 3.1. The code for running the analysis and generating the figures can be found at https://github.com/Larsvanderlaan/npcausalML/blob/main/reproduceSims/introCATEFigure.R.

```
set.seed(12345)
n = 1500
W1 <- rt(n, df = 5)
pi0 <- plogis(W1)
A <- rbinom(n, 1, pi0)
```

```
mu0 =  0.35 + 0.65*plogis(W1-2 )
mu1 <- mu0 + plogis(2*W1+2) - plogis(W1-2) - 0.349
mu <- ifelse(A==1, mu1, mu0)
Y <- rbinom(n, 1, mu)
```

The following `R` code was used to generate the example dataset of Figure 3.2.

```
set.seed(123456)
n = 5
W1 <- rt(n, df = 5)
pi0 <- plogis(W1)
A <- rbinom(n, 1, pi0)
pi <- ifelse(A==1, pi0, 1-pi0)
mu0 =  0.35 + 0.65*plogis( W1-2 )
mu1 <- mu0 + plogis(2*W1+2) - plogis(W1-2) - 0.349
mu <- ifelse(A==1, mu1, mu0)
Y <- rbinom(n, 1, mu)
theta <- W1
weights <- round( (mu1 + mu0 + 1/pi*(Y - mu)),2)
outcome <- round((mu1  + A/pi0*(Y - mu1)) / weights,2)
dat <- data.frame(Covariate = round(W1,3), Weight  = weights, Outcome = outcome)
dat
```

3.B.2 Simulation design

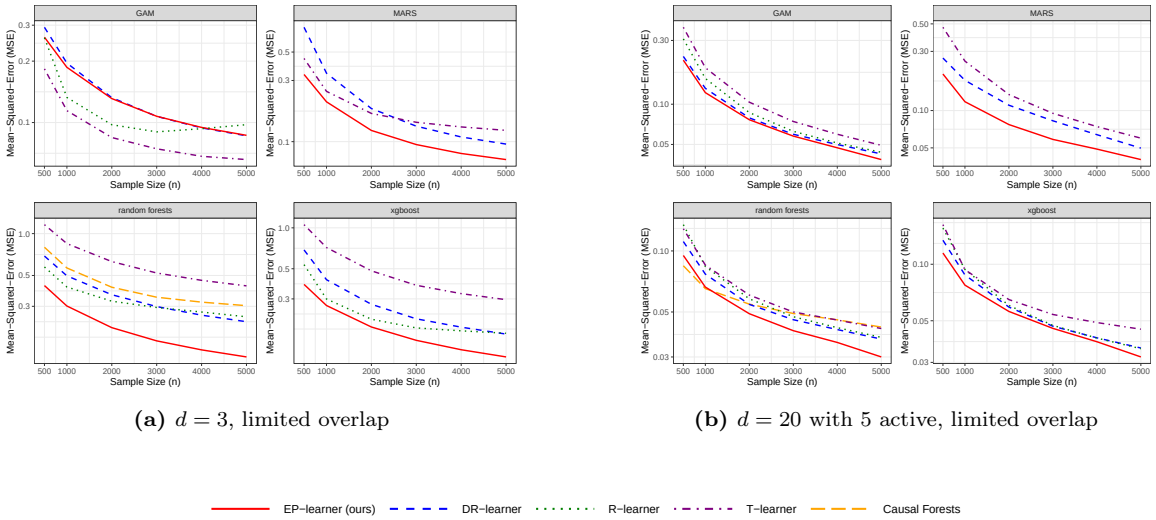
For the low-dimensional CATE experiments, we consider $O = (W_1, W_2, W_3, A, Y)$ for independent covariates W_1, W_2, W_3 each drawn from the uniform distribution on $(-1, +1)$. Given $W = w := (w_1, w_2, w_3)$, the treatment assignment A was generated from a Bernoulli distribution with conditional mean defined, for the moderate overlap setting, as $\pi_0(1 \mid w) := \operatorname{expit}\{(w_1 + w_2 + w_3)/3\}$ and, for the limited overlap setting, as $\pi_0(1 \mid w) := \operatorname{expit}\{w_1 + w_2 + w_3\}$. Given $(W, A) = (w, a)$, the outcome variable Y was generated from a normal distribution with conditional mean $\mu_0(0, w) + a \cdot \theta_0^-(w)$ and variance 4 where $\mu_0(0, w) := \sum_{k=1}^{3}\{w_k/2 + \sin(5w_k) + 1/(w_k + 1.2)\}$ and, in the simple setting, $\theta_0^-(w) := 1 + \sum_{k=1}^{3} w_k$ and, in the complex setting, $\theta_0^-(w) := 1 + \sum_{k=1}^{3}\{w_k + \sin(5w_k)\}$.

For the high-dimensional CATE experiments, we generate $O = (W, A, Y)$ as follows. The covariate W is drawn such that $W/2$ is distributed as a mean-zero multivariate truncated

normal random variable with support $[-2, 2]^{20}$, variances 1, and covariances 0.4. Given $W = w$, the treatment assignment A was generated from a Bernoulli distribution with conditional mean $\pi_0(1 \mid w)$ defined by $\text{logit}\{\pi_0(1 \mid w)\} := (1/1.3)(w_1 + w_5 + w_9 + w_{11} + w_{19})$. Given $(W, A) = (w, a)$, the outcome variable Y was generated from a normal distribution with conditional mean $\mu_0(0, w) + a \cdot \theta_0^-(w)$ and variance 4 where $\mu_0(0, w) = (\cos(4w_1) + \cos(4w_5) + \sin(4w_9) + 1/(1.5 + w_{15}) + 1/(1.5 + w_{10}))/5$ and, in the simple setting, $\theta_0^-(w) := 1 + (w_1 + w_5 + w_9 + w_{15} + w_{10})/5$ and, in the complex setting, $\theta_0^-(w) := 1 + (\sin(4w_1) + \sin(4w_5) + \cos(4w_9) + 1.5(w_{15}^2 - w_{10}^2))/5$.

For the CRR experiments, we consider $O = (W_1, W_2, W_3, A, Y)$ for independent covariates W_1, W_2, W_3 each drawn from the uniform distribution on $(-1, +1)$. Given $W = w$, the treatment assignment A was generated from a Bernoulli distribution with conditional mean $\pi_0(1 \mid w)$ where, in the moderate overlap setting, $\pi_0(1 \mid w) := \text{expit}\{(w_1 + w_2 + w_3)/3\}$ and, in the limited overlap setting, $\pi_0(1 \mid w) := \text{expit}\{w_1 + w_2 + w_3\}$. Given $(W, A) = (w, a)$, the outcome variable Y was generated from a Bernoulli distribution with conditional mean $\mu_0(0, w) \exp\{a\theta_0^{\div}(w)\}$ where $\text{logit}\, \mu_0(0, w) := -1 + 0.3 \sum_{k=1}^{3}\{w_k + \sin(4w_k)\}$ and, in the simple setting, $\theta_0^{\div}(w) := -0.1 + 0.1 \sum_{k=1}^{3} w_k$ and, in the complex setting, $\theta_0^{\div}(w) := -0.1 + 0.1 \sum_{k=1}^{3}\{w_k + \sin(4w_k)\}$.

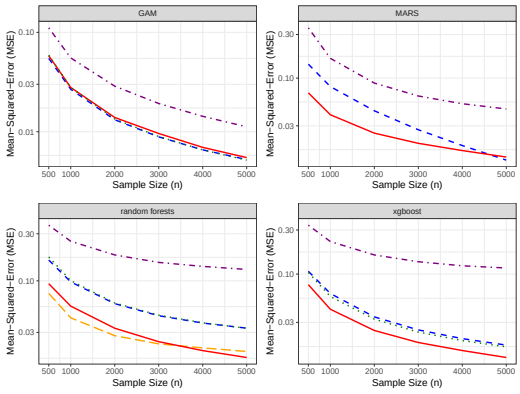
3.B.3 Additional figures



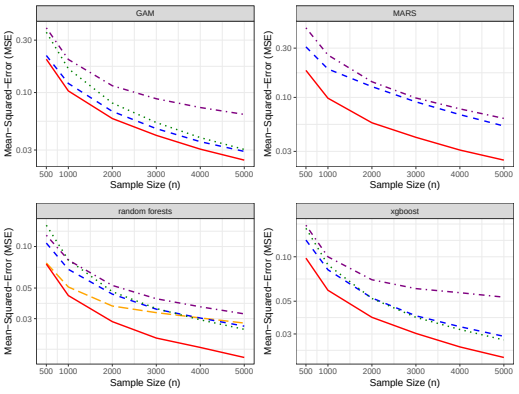
(a) $d = 3$, limited overlap

(b) $d = 20$ with 5 active, limited overlap

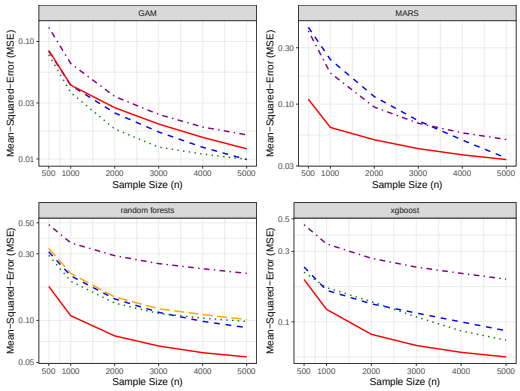
Figure 3.7: CATE experiments with a complex CATE and limited treatment overlap: Mean-squared error for DR-learner and CV-EP-learner with supervised learning algorithms GAM, MARS, ranger, and xgboost. For the plots reporting the results of tree-based algorithms (ranger and xgboost), we also display the results of causal forests for comparison. As a common benchmark across all learners, we display a cross-validated T-learner obtained from an ensemble library including xgboost and GAM learners.



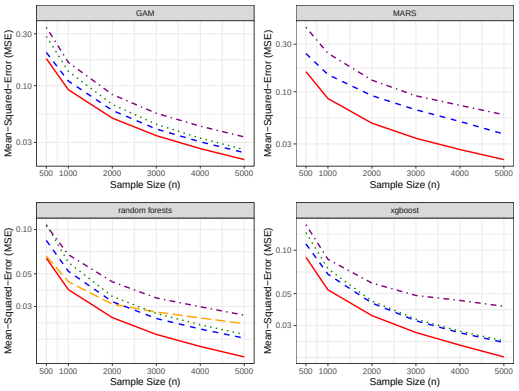
(a) $d = 3$, moderate overlap



(b) $d = 20$ with 5 active, moderate overlap

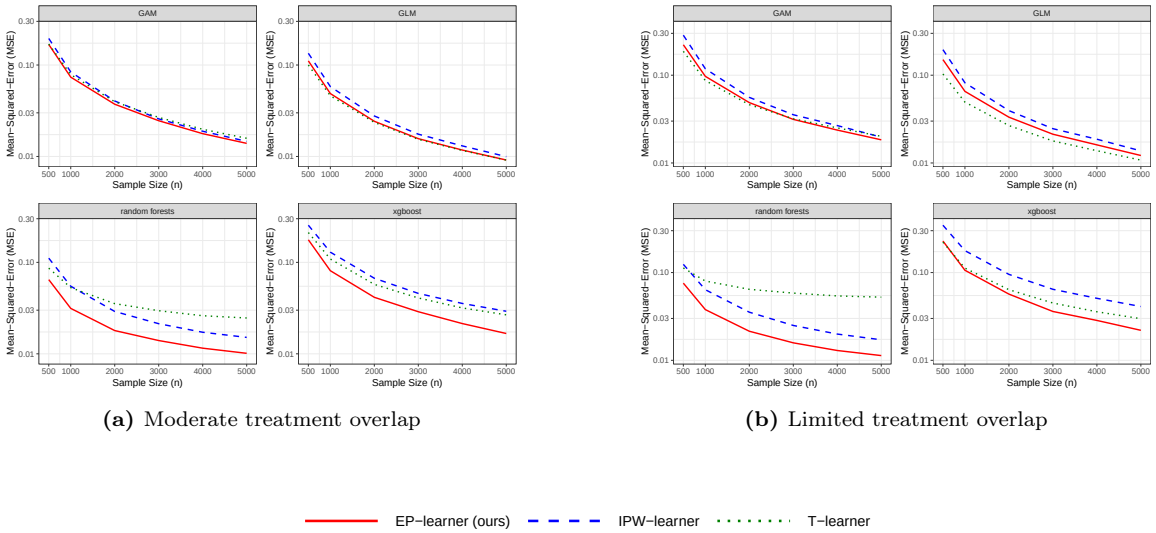


(c) $d = 3$, limited overlap



(d) $d = 20$ with 5 active, limited overlap

Figure 3.8: CATE experiments with a simple CATE and moderate (top) and limited (bottom) treatment overlap: Mean-squared error for DR-learner and CV-EP-learner with supervised learning algorithms GAM, MARS, ranger, and xgboost. For the plots reporting the results of tree-based algorithms (ranger and xgboost), we also display the results of causal forests for comparison. As a common benchmark across all learners, we display a cross-validated T-learner obtained from an ensemble library including xgboost and GAM learners.



(a) Moderate treatment overlap

(b) Limited treatment overlap

Figure 3.9: CRR experiments with a simple CRR and moderate (left) and limited (right) treatment overlap: Mean-squared error for IPW-learner, T-learner, and CV-EP-learner with supervised learning algorithms GAM, random forests, and xgboost. The DR-learner algorithm for the CRR is not implemented due to nonconvexity of the loss.

3.C Additional details on EP-learner algorithm

3.C.1 Simplification of Algorithm 4 for the CATE

When $h_1 \circ \theta = h_2 \circ \theta =: h \circ \theta$ for all $\theta \in \mathcal{F}$, the debiasing term $\frac{1}{n}\sum_{i=1}^{n} \Delta_{\pi_n,\mu_n}(O_i;\theta)$ in the general case simplifies to

$$\frac{1}{n}\sum_{i=1}^{n} \frac{1}{\pi_{n,j(i)}(A_i \,|\, W_i)} \left\{ \sum_{m\in\{1,2\}} H_{m,\mu_{n,j(i)}}(A_i, W_i) \right\} (h \circ \theta)(W_i) \left\{ Y_i - \mu_{n,j(i)}(A_i, W_i) \right\}.$$

In this setting, (3.6) is sufficient to make the debiasing term small. We can instead take the data-dependent feature vector in Algorithm 4 to be

$$\widehat{\varphi}_{k,j(i)}(a,w) = \left\{ \sum_{m\in\{1,2\}} H_{m,\mu_{n,j(i)}}(a,w) \right\} \varphi_k(w).$$

With this modification, the resulting estimator $\mu^*_{n,k(n)}$ will, for all $\psi \in \mathcal{H}_{k(n)}$, satisfy

$$\frac{1}{n}\sum_{i=1}^{n} \frac{1}{\pi_{n,j(i)}(A_i \,|\, W_i)} \left\{ \sum_{m\in\{1,2\}} H_{m,\mu_{n,j(i)}}(A_i, W_i) \right\} \psi(W_i) \left\{ Y_i - \mu^*_{n,j(i)}(A_i, W_i) \right\} = 0.$$

This modification halves the dimension of the sieve regressions of Algorithm 4 without any cost in the size of the debiasing term, and may lead to better finite-sample performance. In particular, for the CATE example, we recommend using $\widehat{\varphi}_{k,j(i)}(a, w) = (2a-1)\varphi_k(w)$, noting that $h_1 \circ \theta = h_2 \circ \theta = \theta$ and $\sum_{m\in\{1,2\}} H_{m,\mu_{n,j(i)}}(a, w) = 2a - 1$.

3.C.2 Debiasing method for general outcomes

Method 3: general outcomes (bound-preserving). Let $\epsilon_n > 0$ be deterministic and define

$$\widehat{a}_\epsilon := \min\left\{ Y_i, \mu_{n,j(i)}(A_i, W_i) : i = 1, \ldots, n \right\} - \epsilon_n,$$
$$\widehat{b}_\epsilon := \max\left\{ Y_i, \mu_{n,j(i)}(A_i, W_i) : i = 1, \ldots, n \right\} + \epsilon_n.$$

Set $g(x) = \mathrm{logit}\{(x - \widehat{a}_\epsilon)/(\widehat{b}_\epsilon - \widehat{a}_\epsilon)\}$. For $i \in [n]$, define

$$\widetilde{Y}_i := \frac{Y_i - \widehat{a}_\epsilon}{\widehat{b}_\epsilon - \widehat{a}_\epsilon}, \qquad \widetilde{\mu}_i := \frac{\mu_{n,j(i)}(A_i, W_i) - \widehat{a}_\epsilon}{\widehat{b}_\epsilon - \widehat{a}_\epsilon}.$$

Let β_n be the coefficient vector obtained from the logistic regression of $\widetilde{Y}_i$ on $\widehat{\varphi}_{k,j(i)}(A_i, W_i)$ with offsets $\mathrm{logit}(\widetilde{\mu}_i)$ and weights $1/\pi_{n,j(i)}(A_i \,|\, W_i)$ for $i \in [n]$. By construction, the sieve-adjusted estimates $\mu^*_{n,j(i)}(A_i, W_i)$ lie in $[\widehat{a}_\epsilon, \widehat{b}_\epsilon]$.

Chapter 4

AUTOMATIC DEBIASED MACHINE LEARNING FOR FUNCTIONALS OF NONPARAMETRIC M-ESTIMANDS

Lars van der Laan, Marco Carone, Aurelien Bibaut, Nathan Kallus, and Alex Luedtke

Abstract

Automatic debiased machine learning (autoDML) reduces valid inference for smooth functionals of regression functions to Riesz representer estimation. We extend this principle to smooth functionals of nonparametric M-estimands, defined as population risk minimizers over infinite-dimensional linear spaces. This formulation covers counterfactual regression functions in static and longitudinal settings, quantile and survival functions, pseudo-outcome regressions such as conditional treatment effects, and solutions to ill-posed inverse problems. The central object is the Riesz representer of the target derivative under the Hessian inner product induced by the population risk. This representer characterizes the leading bias of plug-in estimators and yields the nonparametric efficient influence function for the functional. Debiasing is therefore reduced to two risk minimization problems: one for the M-estimand and one for its Riesz representer. Given their estimators, the debiasing term is obtained by evaluating the loss gradient in the representer direction. We develop the associated first-order theory, including a functional von Mises expansion and the efficiency bound for the resulting projection parameters. The framework accommodates data-driven nuisance estimation through Neyman-orthogonal losses and vector-valued M-estimands defined by coupled objectives. We construct efficient one-step, targeted minimum loss-based, and sieve plug-in estimators, and show that, for quadratic risks and linear functionals, these estimators are asymptotically linear under doubly robust product-rate conditions.

4.1 Introduction

Inference on real-valued functionals of a probability distribution is central to many scientific applications, including treatment-effect estimation, survival analysis, and policy learning. These functionals are often low-dimensional summaries of high- or infinite-dimensional nuisance functions [Bickel et al., 1993, van der Laan and Robins, 2003, van der Laan and Rose, 2011, Chernozhukov et al., 2018a]. A canonical example is the average treatment effect under full confounding adjustment, which can be written as the difference between two partial means of the outcome regression function [Bang and Robins, 2005]. More generally, nuisance

functions are often characterized as minimizers of population risks, such as mean-squared error in regression or negative log-likelihood in survival analysis. The flexibility of modern machine learning makes it attractive for estimating such nuisance functions. However, simply applying the target functional to a machine-learned nuisance estimate can introduce substantial bias, converge too slowly for valid inference, and does not by itself provide confidence intervals or other inferential guarantees.

Several methods have been developed to correct the bias of naive plug-in estimators and obtain $\sqrt{n}$-consistent, asymptotically normal inference under suitable conditions. These include classical one-step estimators [Levit, 1975, Pfanzagl and Wefelmeyer, 1985, Hasminskii and Ibragimov, 1979, Bickel et al., 1993, Newey, 1994, Newey et al., 1998], estimating-equation and double/debiased machine learning approaches [Robins et al., 1995, 2000, van der Laan and Robins, 2003, Chernozhukov et al., 2018a], targeted minimum loss-based estimation [Laan and Rubin, 2006, van der Laan and Rose, 2011], and sieve-based estimation strategies [Qiu et al., 2021, van der Laan and Rose, 2021, van der Laan et al., 2022, Cho et al., 2024, Li et al., 2025]. These methods share a common principle: estimate the nuisance functions flexibly, then add a correction that removes the leading plug-in bias. The main difficulty is that this correction is usually parameter-specific. Constructing it requires deriving the efficient influence function, often using differential-geometric and functional-analytic tools, and identifying and estimating any additional nuisance functions needed for asymptotically linear inference [Bickel et al., 1993]. Thus, while debiasing provides a general route to valid inference with flexible nuisance estimation, its implementation remains technically demanding and largely case-specific.

A growing literature seeks to automate parts of this construction in semiparametric statistics and causal inference. One line of work approximates efficient influence functions numerically for pathwise differentiable functionals [Frangakis et al., 2015, Luedtke et al., 2015, Carone et al., 2019, Ichimura and Newey, 2022, Jordan et al., 2022b]. These methods are broadly applicable, but can be computationally demanding and tuning-sensitive because they perturb the data distribution and numerically differentiate the parameter separately for each observation. A second line of work uses automatic differentiation [Rall, 1981] to construct influence functions for parameters built from primitives with known Hilbert-valued pathwise derivatives [Luedtke and Chung, 2024, Luedtke, 2024]. This approach automates the assembly of influence functions for composite parameters, but relies on a library of primitives whose derivatives have already been derived.

A complementary approach, known as *automatic debiased machine learning* (autoDML), exploits problem structure to derive reusable influence-function representations [Ichimura and Newey, 2022, Chernozhukov et al., 2022c, Rotnitzky et al., 2021]. In nonparametric models, these representations are efficient influence functions; under additional model restrictions, they may instead be valid influence functions for the corresponding restricted model or projection target. For linear functionals of regression functions, the representation is determined by one additional nuisance: the Riesz representer of the target functional [Newey, 1994, Chernozhukov et al., 2021b, 2022c]. When this representer is not available analytically, it can be learned through a second risk minimization problem, often called *Riesz regression* [Chernozhukov et al., 2021b, 2022d, 2026, Williams et al., 2025, Hines and Miles, 2025]. This perspective makes debiasing modular: once the influence-function form is

known, the remaining task is to estimate the required nuisance components, including any representers, using flexible regression methods [Breiman, 2001, Freund and Schapire, 1997, Abdi et al., 1999, Chernozhukov et al., 2021b, Lee and Schuler, 2025].

Existing autoDML remains anchored in regression settings, where the target is typically a functional of a real-valued regression function defined as the minimizer of a known loss. Important extensions cover covariate shift [Chernozhukov et al., 2023], sequential regression problems [Chernozhukov et al., 2022c], and linear inverse problems with regression-type nuisance functions [Bennett et al., 2023b, 2025, van der Laan et al., 2025c]. Although these extensions share a common logic, they are developed separately, with setting-specific representers, nuisance components, and debiasing arguments. Thus, existing autoDML provides powerful procedures for important problem classes, but not yet a unified account of debiasing smooth functionals of nonparametric M-estimands, defined as population risk minimizers over infinite-dimensional linear spaces.

Many parameters of current interest can be expressed as functionals of M-estimands. Examples include functionals of counterfactual regression functions [Rubin and van der Laan, 2007], value functions in reinforcement learning [van der Laan et al., 2025c], density ratios [Hines and Miles, 2025], quantile and survival functions [Wang et al., 2018, Westling et al., 2024], pseudo-outcome regressions such as conditional average treatment effects [Yang et al., 2023, Foster and Syrgkanis, 2023], and smooth functionals of solutions to potentially ill-posed inverse problems [Carrasco et al., 2007, Bennett et al., 2025, van der Laan et al., 2025c]. This raises the central question of this paper: beyond regression, what common structure determines the influence function and debiasing correction for smooth functionals of M-estimands?

4.1.1 Contributions of this work

We give a unified construction of efficient influence functions, debiasing corrections, and autoDML estimators for smooth functionals of M-estimands defined by general risk functions in nonparametric statistical models. The key object is a single auxiliary function: the Riesz representer of the target derivative under the Hessian inner product induced by the local curvature of the risk. When the loss is nuisance-free, or Neyman-orthogonal with respect to its nuisance components, this representer determines both the nonparametric efficient influence function and the first-order debiasing correction. The correction is obtained by evaluating the loss derivative in the corresponding Riesz direction.

This characterization makes the construction modular. One differentiates the target and the loss, computes the representer analytically or through an auxiliary risk minimization problem, and substitutes it into the general formula. Thus, the framework provides a systematic route to efficient influence-function derivations and automatic debiased estimators for parameters defined through nuisance-dependent losses, nonlinear targets, and coupled vector-valued M-estimands.

Our main contributions are as follows:

(i) We formalize smooth functionals of M-estimands as *nonparametric projection parameters* and derive a modular characterization of their von Mises expansion, efficient influence function, and efficiency bound.

(ii) We develop a general autoDML framework for these parameters, including one-step, targeted minimum loss-based, and sieve plug-in estimators.

(iii) We establish asymptotic linearity and nonparametric efficiency of the proposed estimators. For quadratic risks and linear functionals, the resulting estimators are doubly robust in the M-estimand and Riesz representer.

4.1.2 Related work

This paper contributes to the literature on automatic debiased machine learning for low-dimensional functionals of infinite-dimensional parameters [Chernozhukov et al., 2022c, 2021b, 2023, Bennett et al., 2023b, 2025]. Chernozhukov et al. [2022c] study functionals of outcome regressions, and Chernozhukov et al. [2021b] extend this approach to generalized linear models. Related influence-function representations have been derived for functionals whose nuisance components are characterized by exogenous or endogenous orthogonality conditions [Ichimura and Newey, 2022], general conditional moment models [Argañaraz, 2025], and parameters with mixed-bias remainders [Rotnitzky et al., 2021, 2026]. In missing-data and causal inference problems, related work shows how incomplete-data influence functions can be obtained from complete-data influence functions through augmentation [Robins et al., 1994, 1995, 2000, van der Laan and Robins, 2003, Rose and van der Laan, 2011, Graham et al., 2024]. Across these settings, suitable Riesz representers provide a common route to debiasing. These works, however, are developed around regression, moment restrictions, or incomplete-data reductions, and do not give a unified loss-based theory for smooth functionals of general M-estimands.

The closest related work is Section 3 of Chernozhukov et al. [2024a], which studies functionals of real-valued M-estimands defined as minimizers of known losses, as well as separable collections of such estimands. Although motivated by M-estimation, that work follows the orthogonal-moment approach of Ichimura and Newey [2022] and expresses the analysis through generalized residuals rather than through the Hessian geometry of the population risk. It does not allow the loss to depend on unknown nuisance components, nor does it cover vector-valued M-estimands defined jointly by a coupled objective. Our framework works directly with the loss-based structure and extends automatic debiasing to these settings.

We place these problems in a semiparametric framework based on nonparametric projection parameters. This yields a functional von Mises expansion, the efficient influence function, and the efficiency bound for smooth functionals of general M-estimands. To our knowledge, such an efficiency theory has not previously been developed in the autoDML literature. The central technical object is the Riesz representer under the Hessian inner product induced by the population risk. Thus, automatic debiasing is determined jointly by the local derivative of the target functional and the local curvature of the risk. In the setting of Chernozhukov et al. [2024a], this object reduces to their Riesz representer.

We also extend the estimator side of autoDML. In addition to the one-step estimators emphasized in prior work, we study automatic targeted minimum loss-based estimators and automatic sieve plug-in estimators within the same framework. The automatic TMLE

updates the M-estimand itself to obtain an efficient plug-in estimator, while the sieve construction provides an automatic undersmoothing rule for selecting a sieve dimension large enough to approximate both the M-estimand and its Riesz representer.

Our results also connect autoDML to the classical literature on sieve M-estimation, where analogous representers appear in proofs of asymptotic normality for plug-in estimators [Shen, 1997, Chen et al., 2014, Chen and Liao, 2014, Chen, 2007, Qiu et al., 2021]. In our framework, these representers are the directions that determine the efficient influence-function correction for the corresponding nonparametric projection parameter. Consequently, sieve plug-in estimators can be viewed as one-step estimators with an implicit debiasing term, and are efficient for the projection parameter induced by the chosen loss. This efficiency is loss-specific: it need not coincide with the sharp efficiency bound in a restricted model where the sieve space is imposed as correctly specified. For example, when the loss differs from the log-likelihood loss, the resulting estimator may fail to attain the restricted-model efficiency bound, recovering the classical inefficiency of least-squares sieve estimators under heteroscedasticity [Shen, 1997]. Our framework also extends sieve-style inference beyond known losses. Existing two-stage sieve methods allow nuisance-dependent losses, but typically in nonorthogonal settings where the nuisance components are themselves estimated by sieves [Hahn et al., 2018]. By contrast, Neyman-orthogonal losses separate nuisance estimation from the target M-estimation problem, allowing the nuisance components to be estimated by generic machine-learning methods.

Organization. This paper is organized as follows. Section 4.2 introduces the framework, presents illustrative examples, and reviews related literature. Section 4.3 derives the functional von Mises expansion and the efficient influence function for functionals of M-estimands. Section 4.4 presents the proposed estimators, and Section 4.5 develops their asymptotic theory. Section 4.6 illustrates the framework through an application to long-term survival in the beta-geometric model.

4.2 Generalized autoDML framework

4.2.1 Problem setup

Suppose $Z_1, \ldots, Z_n$ are i.i.d. draws from P_0, where P_0 belongs to a convex nonparametric model $\mathcal{P}$. We consider parameters of the form

$$\Psi(P_0) = \psi_{P_0}(\theta_{P_0}) = E_{P_0}\{m(Z, \theta_{P_0})\},$$

where $m : \mathcal{Z} \times \mathcal{H} \to \mathbb{R}$ is known and θ_{P_0} is an M-estimand. More generally, for each $P \in \mathcal{P}$, θ_P takes values in a normed linear space $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ and is the unique minimizer

$$\theta_P = \operatorname*{argmin}_{\theta \in \mathcal{H}} L_P(\theta, \eta_P). \tag{4.1}$$

Here $L_P(\theta, \eta) := E_P\{\ell_\eta(Z, \theta)\}$ is the risk induced by a known loss $\ell_\eta : \mathcal{Z} \times \mathcal{H} \to \mathbb{R}$. The loss may depend on a nuisance component η, which takes values in a convex subset $\mathcal{N}$ of an ambient normed space $(\widetilde{\mathcal{N}}, \|\cdot\|_{\mathcal{N}})$; η_P denotes the corresponding nuisance under P. Thus, for any $P \in \mathcal{P}$, $\Psi(P) = \psi_P(\theta_P)$, where $\psi_P(\theta) := E_P\{m(Z, \theta)\}$. We write S_0 for S_{P_0}; for example, $\eta_0 = \eta_{P_0}$, $\theta_0 = \theta_{P_0}$, and $\psi_0 = \psi_{P_0}$. We also write $Pf := \int f(z)\, dP(z)$.

We assume that the population risk $L_0(\theta, \eta)$ is sufficiently smooth near (θ_0, η_0): it is twice Fréchet differentiable in θ and once differentiable with respect to nuisance perturbations. For $h, h_1, h_2 \in \mathcal{H}$ and $\eta \in \mathcal{N}$, write

$$\begin{aligned}
\partial_\theta L_0(\theta_0, \eta_0)(h) &:= \frac{d}{dt} L_0(\theta_0 + th, \eta_0)\Big|_{t=0}, \\
\partial_\theta^2 L_0(\theta_0, \eta_0)(h_1, h_2) &:= \frac{\partial^2}{\partial t\, \partial s} L_0(\theta_0 + th_1 + sh_2, \eta_0)\Big|_{t=0,s=0}, \\
\partial_\eta \partial_\theta L_0(\theta_0, \eta_0)(\eta - \eta_0, h) &:= \frac{\partial^2}{\partial t\, \partial s} L_0(\theta_0 + th, \eta_0 + s(\eta - \eta_0))\Big|_{t=0,s=0}.
\end{aligned}$$

We also assume that $\ell_{\eta_0}(\cdot, \theta)$ is Gâteaux differentiable in θ, with derivative $\dot{\ell}_{\eta_0}(\theta_0) : \mathcal{H} \to L^2(P_0)$ satisfying

$$\partial_\theta L_0(\theta_0, \eta_0)(h) = P_0 \dot{\ell}_{\eta_0}(\theta_0)(h).$$

Finally, we assume that the Hessian bilinear form $(h_1, h_2) \mapsto \partial_\theta^2 L_0(\theta_0, \eta_0)(h_1, h_2)$ is positive definite, and hence defines an inner product on $\mathcal{H}$. This inner product and the loss derivative are the primitive objects used below to construct the debiasing correction. Section 4.3 states the full smoothness assumptions.

Our framework also allows the loss to depend on an unknown nuisance component η_0. We assume that the loss has already been orthogonalized with respect to this nuisance, so that, at (θ_0, η_0), for all $\eta \in \mathcal{N}$ and $h \in \mathcal{H}$,

$$\partial_\eta \partial_\theta L_0(\theta_0, \eta_0)(\eta - \eta_0, h) = 0.$$

This is an assumption on the loss representation, not a substantive restriction on the statistical problem. It makes the first-order analysis equivalent to the known-nuisance case, while still allowing η_0 to be estimated flexibly. In many settings, such representations can be obtained from nonorthogonal formulations by standard orthogonalization arguments [Rubin and van der Laan, 2007, Yang et al., 2023, Foster and Syrgkanis, 2023].

Remark 4.1 (Separable product spaces and vector-valued M-estimands)**.** Our framework supports functionals of multiple M-estimands that jointly minimize a common objective. A simple special case is when each M-estimand solves a separate objective. Suppose that $\mathcal{H} = \mathcal{H}_1 \times \cdots \times \mathcal{H}_K$, where each $\mathcal{H}_k$ is a normed space, and write $\theta = (\theta_1, \ldots, \theta_K)$. Consider a loss of the form $\ell_{\eta_0}(z, \theta) = \sum_{k=1}^K \ell_{\eta_0}^{(k)}(z, \theta_k)$. Then the population risk is separable: $L_0(\theta, \eta_0) = \sum_{k=1}^K L_0^{(k)}(\theta_k, \eta_0)$, where $L_0^{(k)}(\theta_k, \eta_0) := P_0 \ell_{\eta_0}^{(k)}(\cdot, \theta_k)$. Thus, each component $\theta_{0,k}$ minimizes its own risk $L_0^{(k)}(\cdot, \eta_0)$. Moreover, for $h = (h_1, \ldots, h_K)$ and $h' = (h'_1, \ldots, h'_K)$, $\dot{\ell}_{\eta_0}(\theta_0)(h) = \sum_{k=1}^K \dot{\ell}_{\eta_0}^{(k)}(\theta_{0,k})(h_k)$ and $\partial_\theta^2 L_0(\theta_0, \eta_0)(h, h') = \sum_{k=1}^K \partial_{\theta_k}^2 L_0^{(k)}(\theta_{0,k}, \eta_0)(h_k, h'_k)$.

4.2.2 Estimator and overview of theory

Let η_n and θ_n be estimators of the nuisance function η_0 and the M-estimand θ_0. For example, θ_n may be obtained by empirical risk minimization based on the orthogonal loss $\theta \mapsto n^{-1} \sum_{i=1}^n \ell_{\eta_n}(Z_i, \theta)$. A natural estimator of $\Psi(P_0)$ is the plug-in estimator

$$\psi_n(\theta_n) := \frac{1}{n} \sum_{i=1}^n m(Z_i, \theta_n).$$

Yet when θ_n is estimated using flexible learning methods, $\psi_n(\theta_n)$ typically does not admit $n^{1/2}$-rate inference. The difficulty is that the target depends on θ_n at first order, so estimation error in $\theta_n - \theta_0$ propagates directly into the plug-in estimator [van der Laan and Rose, 2011, Chernozhukov et al., 2018a], even when η_0 is known. Debiasing is therefore needed to remove this first-order sensitivity and restore asymptotic linearity.

For linear functionals of a regression function $\theta_0 : x \mapsto E_0[Y \mid X = x]$, with $Z = (X, Y)$, Chernozhukov et al. [2022c] developed an automatic approach to debiasing the plug-in estimator $\psi_n(\theta_n)$. By the Riesz representation theorem, they showed that the plug-in error $\psi_n(\theta_n) - \psi_n(\theta_0)$ can be written as $\langle \alpha_0, \theta_n - \theta_0 \rangle_{L^2(P_{0,X})} = \int \alpha_0(x)\{\theta_n(x) - \theta_0(x)\}\, dP_{0,X}(x) = \int \alpha_0(x)\{\theta_n(x) - y\}\, dP_0(z)$ for an appropriate representer α_0. This yields the automatic DML estimator

$$\psi_n(\theta_n) - \frac{1}{n}\sum_{i=1}^{n} \alpha_n(X_i)\{\theta_n(X_i) - Y_i\}.$$

The second term corrects the plug-in error through a one-step bias correction [Bickel et al., 1993], thereby enabling fast rates and valid inference for $\Psi(P_0)$, even when α_0 and θ_0 are estimated at slow rates. This estimator generalizes the well-known augmented inverse probability weighted estimator [Robins et al., 2000, 1995, Bruns-Smith et al., 2025]. A key observation of Chernozhukov et al. [2022c] is that the Riesz representer α_0 can itself be characterized as the solution to a risk minimization problem:

$$\arg\min_{\alpha \in \mathcal{H}} E_0\{\alpha(X)^2 - 2\, m(Z, \alpha)\} = \arg\min_{\alpha \in \mathcal{H}} E_0\big\{\alpha(X)^2 - 2\alpha(X)\alpha_0(X)\big\} = \arg\min_{\alpha \in \mathcal{H}} E_0\big[\{\alpha(X) - \alpha_0(X)\}^2\big].$$

This follows from the defining property of the Riesz representer, namely $E_0\{m(Z, \alpha)\} = E_0\{\alpha(X)\alpha_0(X)\}$ [Williams et al., 2025]. Based on this characterization, they develop autoDML estimators by estimating θ_0 using the squared-error loss $\ell_\eta : (z, \theta) \mapsto \{y - \theta(x)\}^2$, and estimating α_0 via *Riesz regression*, that is, by minimizing the Riesz loss $(z, \alpha) \mapsto \alpha(x)^2 - 2\, m(z, \alpha)$.

To extend this idea beyond regression functionals, it is useful to recall the classical finite-dimensional M-estimation setting. Let $\theta_0 = \operatorname{argmin}_{\theta \in \mathbb{R}^d} P_0 \ell_{\eta_0}(\cdot, \theta)$. Given preliminary estimators θ_n and η_n, the classical one-step estimator [van der Vaart, 2000] is

$$\widehat{\theta}_n^{\mathrm{os}} := \theta_n - H_n^{-1} P_n \nabla \ell_{\eta_n}(\cdot, \theta_n), \qquad H_n := P_n \nabla^2 \ell_{\eta_n}(\cdot, \theta_n).$$

For a linear target $\psi(\theta) = b^\top \theta$, set $\alpha_n := H_n^{-1} b$. Then

$$\psi(\widehat{\theta}_n^{\mathrm{os}}) = \psi(\theta_n) - P_n \dot{\ell}_{\eta_n}(\theta_n)(\alpha_n), \qquad \dot{\ell}_{\eta_n}(\theta_n)(\alpha) := \alpha^\top \nabla \ell_{\eta_n}(\cdot, \theta_n).$$

Thus the one-step correction is the empirical directional derivative of the loss in the direction α_n. This direction is characterized by

$$\psi(\theta) = \partial_\theta^2 L_n(\theta_n, \eta_n)(\theta, \alpha_n) \qquad \text{for all } \theta \in \mathbb{R}^d,$$

where $L_n(\theta, \eta) := P_n \ell_\eta(\cdot, \theta)$. Hence, α_n is the empirical Riesz representer of ψ under the Hessian-induced bilinear form. For squared-error loss, $\ell(z, \theta) = \{y - \theta\}^2/2$, this form is Euclidean, so $\alpha_n = b$ and $\dot{\ell}(\theta_n)(\alpha_n) = \alpha_n^\top(\theta_n - y)$, recovering the correction in Chernozhukov et al. [2022c]. This finite-dimensional case is the prototype for the construction below.

We now extend this idea to smooth functionals of M-estimands. For a possibly nonlinear target ψ_0, the relevant first-order object is its Fréchet derivative $\dot{\psi}_0(\theta_0) : \mathcal{H} \to \mathbb{R}$. The finite-dimensional calculation suggests defining the debiasing direction as the Hessian Riesz

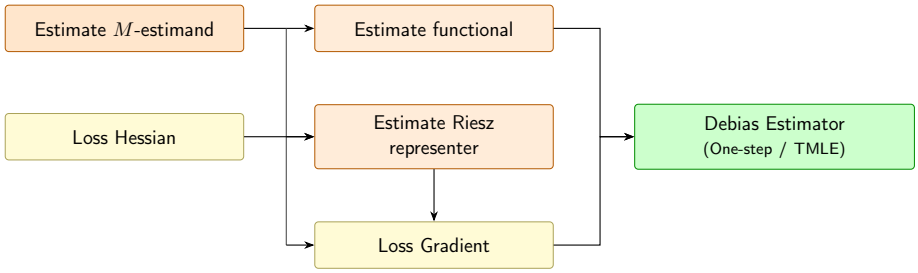


Figure 4.1: Given a specification of the orthogonal loss function and the target functional of the M-estimand, the diagram illustrates the autoDML workflow.

representer $\alpha_0 \in \overline{\mathcal{H}}$, for a suitable completion $\overline{\mathcal{H}}$ of $\mathcal{H}$, satisfying

$$\dot{\psi}_0(\theta_0)(h) = \partial_\theta^2 L_0(\theta_0, \eta_0)(h, \alpha_0), \qquad h \in \mathcal{H}. \tag{4.2}$$

Equivalently, α_0 minimizes the quadratic criterion

$$\alpha_0 = \operatorname*{argmin}_{\alpha \in \overline{\mathcal{H}}} \left\{ \partial_\theta^2 L_0(\theta_0, \eta_0)(\alpha, \alpha) - 2\dot{\psi}_0(\theta_0)(\alpha) \right\}. \tag{4.3}$$

Indeed, using (4.2), the objective in (4.3) equals

$$\partial_\theta^2 L_0(\theta_0, \eta_0)(\alpha - \alpha_0, \alpha - \alpha_0) - \partial_\theta^2 L_0(\theta_0, \eta_0)(\alpha_0, \alpha_0),$$

and is therefore minimized at α_0. Given estimators θ_n, η_n, and α_n, the generalized autoDML estimator is

$$\widehat{\psi}_n := \psi_n(\theta_n) - P_n \dot{\ell}_{\eta_n}(\theta_n)(\alpha_n). \tag{4.4}$$

The construction is automatic in the sense that, once the loss ℓ_η, the target map ψ_0, and the nuisance estimator η_n are specified, inference reduces to estimating θ_0 and its Hessian representer α_0. The latter can often be learned by risk minimization whenever the quadratic criterion in (4.3) admits the pointwise form

$$(z, \alpha) \mapsto \ddot{\ell}_{\eta_0}(\theta_0)(\alpha, \alpha)(z) - 2\, \dot{m}_{\theta_0}(z, \alpha),$$

where $\ddot{\ell}_{\eta_0}(\theta_0)$ and $\dot{m}_{\theta_0}$ are the corresponding second- and first-order directional derivatives. In many applications, these derivatives are already available from the procedure used to estimate θ_0, such as gradient-based optimization, boosting, or classical M-estimation; otherwise, they can often be derived analytically from the loss. Sufficient smoothness conditions are given in Lemma 4.7.

Heuristically, the correction remains valid in infinite-dimensional settings because the leading plug-in bias is governed by the local quadratic geometry of the risk. By a first-order expansion of the target and the Hessian Riesz representation,

$$\begin{aligned} \psi_0(\theta_n) - \psi_0(\theta_0) &\approx \dot{\psi}_0(\theta_0)(\theta_n - \theta_0) \\ &= \partial_\theta^2 L_0(\theta_0, \eta_0)(\theta_n - \theta_0, \alpha_0) \\ &= \partial_\theta^2 L_0(\theta_0, \eta_0)(\theta_n - \theta_0, \alpha_n) + \partial_\theta^2 L_0(\theta_0, \eta_0)(\theta_n - \theta_0, \alpha_0 - \alpha_n). \end{aligned}$$

The last term is controlled by a product rate in the errors of θ_n and α_n. For the leading

term, a Taylor expansion of the risk gradient gives

$$\partial_\theta L_0(\theta_n, \eta_0)(\alpha_n) - \partial_\theta L_0(\theta_0, \eta_0)(\alpha_n) \approx \partial_\theta^2 L_0(\theta_0, \eta_0)(\theta_n - \theta_0, \alpha_n).$$

Since θ_0 minimizes $L_0(\cdot, \eta_0)$, the second term is zero. Neyman orthogonality permits replacing η_0 by η_n to first order, so that

$$\psi_0(\theta_n) - \psi_0(\theta_0) \approx \partial_\theta L_0(\theta_n, \eta_n)(\alpha_n) = P_0 \dot{\ell}_{\eta_n}(\theta_n)(\alpha_n).$$

Thus the empirical directional derivative in (4.4) estimates the leading plug-in bias, and subtracting it yields a first-order debiased estimator. Section 4.3.2 makes this expansion precise.

This heuristic guides our asymptotic analysis in Section 4.5. Under suitable rate conditions on η_n, θ_n, and α_n, we show that

$$\widehat{\psi}_n - \Psi(P_0) = P_n \chi_0 + o_p(n^{-1/2}), \qquad \chi_0(z) := -\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0)(z) + m(z, \theta_0) - \Psi(P_0).$$

Here χ_0 is the efficient influence function of Ψ in the nonparametric model. For linear targets and quadratic risks, the expansion holds under the doubly robust product-rate condition

$$\partial_\theta^2 L_0(\theta_0, \eta_0)(\theta_n - \theta_0, \alpha_n - \alpha_0) = o_p(n^{-1/2}),$$

which extends the familiar double-robustness phenomenon for linear functionals of regression functions [Bang and Robins, 2005, Chernozhukov et al., 2018a]. Consequently, $\widehat{\psi}_n$ is regular and nonparametrically efficient, and $n^{1/2}\{\widehat{\psi}_n - \Psi(P_0)\}$ is asymptotically normal with mean zero and variance $\sigma_0^2 := \mathrm{Var}_0\{\chi_0(Z)\}$. Wald confidence intervals and tests may therefore be based on the plug-in variance estimator

$$\frac{1}{n}\sum_{i=1}^{n} \chi_n(Z_i)^2, \qquad \chi_n(z) := -\dot{\ell}_{\eta_n}(\theta_n)(\alpha_n)(z) + m(z, \theta_n) - \psi_n(\theta_n).$$

4.2.3 Examples

We next show that the framework covers several familiar parameters from semiparametric statistics and causal inference. First, it includes both regression functionals and functionals of Riesz representers, such as counterfactual means and average treatment effects. Second, by allowing losses to depend on nuisance parameters, it covers target-specific M-estimands, such as treatment effect modifiers in semiparametric regression models. Third, it covers smooth functionals of solutions to potentially ill-posed inverse problems, once those solutions are characterized as minimizers of risks induced by Neyman-orthogonal losses. Further details are given in Appendix 4.D.

Throughout the first examples, let $Z = (X, A, Y)$, where $X \in \mathbb{R}^d$, $A \in \{0, 1\}$, and $Y \in \mathbb{R}$, and write $\mu_P(a, x) := E_P[Y \mid A = a, X = x]$.

Example 1a (Functionals of representers and regressions)**.** Let $r : \mathcal{Z} \times \mathcal{H} \to \mathbb{R}$ be linear in its second argument and suppose that, for some $M < \infty$, $|E_0\{r(Z, \theta)\}| \le M\|\theta\|_{L^2(P_0)}$ for all $\theta \in \mathcal{H}$. If θ_0 is the $L^2(P_0)$ Riesz representer of $\theta \mapsto E_0\{r(Z, \theta)\}$, then θ_0 is an M-estimand under the quadratic loss

$$\ell(z, \theta) := \frac{1}{2}\theta(a, x)^2 - r(z, \theta)$$

[Chernozhukov et al., 2022c]. This formulation covers both regression and weighting objects [Williams et al., 2025]. The outcome regression $\mu_0(a, x)$ is obtained by taking $r(z, \theta) =$

$y\theta(a,x)$, whereas the inverse-probability weight $w_0(a,x) := 1(a=1)/P_0(A=1 \mid X=x)$ is obtained by taking $r(z,\theta) = \theta(1,x)$. Thus the counterfactual mean under treatment can be written either as the regression functional $E_0\{\mu_0(1,X)\}$ or as the weighting functional $E_0\{w_0(A,X)Y\}$. For this loss,

$$\dot{\ell}(\theta)(\alpha)(z) = \alpha(a,x)\theta(a,x) - r(z,\alpha),$$

and the Hessian-induced inner product is the usual $L^2(P_{0,X,A})$ inner product. More generally, Riesz losses based on Bregman divergences can be used [Hines and Miles, 2025].

The next example illustrates why it is useful to work with general M-estimands rather than only regression functions. In semiparametric regression models, effect modifiers such as conditional average treatment effects, relative risks, and log-odds differences can be defined directly as minimizers of orthogonal losses [Nekipelov et al., 2022, van der Laan et al., 2024a]. This permits inference on functionals of the effect modifier without first estimating the full outcome regression, and allows $\mathcal{H}$ to encode structure in the target itself. For example, in a partially linear model, the CATE can be estimated by optimizing directly over a finite-dimensional linear class.

Example 1b (Functionals in semiparametric regression models)**.** Let $g : \mathbb{R} \to \mathbb{R}$ be a known link function and suppose that

$$g\{\mu_0(A,X)\} = h_0(X) + A\tau_0(X),$$

where $h_0 := g\{\mu_0(0,\cdot)\}$ is unrestricted and $\tau_0 \in \mathcal{T} \subset L^2(P_{0,X})$. In many applications, the target is a functional $\psi_0(\tau_0)$ of the treatment effect modifier τ_0, rather than a functional of the full outcome regression. Under suitable orthogonal losses, τ_0 is itself an M-estimand [Nekipelov et al., 2022, van der Laan et al., 2024a].For binary outcomes with logistic link, one such loss is

$$\ell_{\eta_0}(z,\theta) := \frac{1}{\nu_0(a,x)}\left[\log\{1+\exp(-\{a-\pi_0(x)\}\theta(x) - h_0(x))\} - y\{a-\pi_0(x)\}\theta(x)\right],$$

where $\eta_0 = (\pi_0, \mu_0)$, $\pi_0(x) := P_0(A=1 \mid X=x)$,

$$h_0(x) := \pi_0(x)\operatorname{logit}\mu_0(1,x) + \{1-\pi_0(x)\}\operatorname{logit}\mu_0(0,x), \qquad \nu_0(a,x) := \mu_0(a,x)\{1-\mu_0(a,x)\}.$$

The corresponding loss for the Hessian Riesz representer α_0 is

$$(z,\alpha) \mapsto \{a-\pi_0(x)\}^2\alpha(x)^2 - 2\,\dot{m}_{\theta_0}(z,\alpha),$$

which can be implemented as a Riesz regression with overlap weights [Li et al., 2019].

Our framework also covers smooth functionals of pseudo-outcome regressions defined by Neyman-orthogonal least-squares losses or conditional influence functions [Rubin and van der Laan, 2007, Kennedy et al., 2017, Kennedy, 2023b, Yang et al., 2023, Chernozhukov et al., 2024b]. Such losses are common in causal inference and missing-data problems, where orthogonalization reduces first-order sensitivity to nuisance estimation. Pseudo-outcome regressions arise, for example, in CATE estimation with instrumental variables [Syrgkanis et al., 2019], causal dose–response estimation [Kennedy et al., 2017, Westling et al., 2020], and proximal causal inference [Yang et al., 2023, Sverdrup and Cui, 2023].

Example 1c (Functionals of pseudo-outcome regressions)**.** Many Neyman-orthogonal losses can be written as

$$\ell_{\eta_0}(z,\theta) = \frac{w_{\eta_0}(z)}{2}\{\zeta_{\eta_0}(z) - \theta(x)\}^2,$$

where w_{η_0} is a pseudo-weight and ζ_{η_0} is a pseudo-outcome. The derivative is $\dot{\ell}_\eta(\theta)(\alpha)(z) = w_\eta(z)\alpha(x)\{\theta(x) - \zeta_\eta(z)\}$, and the Hessian representer solves

$$\alpha_0 = \operatorname*{argmin}_{\alpha \in \mathcal{H}} E_0\left[w_{\eta_0}(Z)\alpha(X)^2 - 2\,\dot{m}_{\theta_0}(Z, \alpha)\right].$$

This form includes the augmented inverse-probability weighted loss for regression with missing outcomes [Rubin and van der Laan, 2007, van der Laan and Dudoit, 2003]. If Y is observed only when $R = 1$, $Z = (X, R, RY)$, $\pi_0(x) := P_0(R = 1 \mid X = x)$, and $\mu_0(x) := E_0[Y \mid R = 1, X = x]$, then under $Y \perp R \mid X$, the full-data regression $\theta_0(x) := E_0[Y \mid X = x]$ is identified by $\mu_0(x)$. The orthogonal squared-error loss is obtained by taking $w_{\eta_0}(z) = 1$ and

$$\zeta_{\eta_0}(z) = \mu_0(x) + \frac{r}{\pi_0(x)}\{y - \mu_0(x)\}, \qquad \eta_0 = (\pi_0, \mu_0).$$

The same pseudo-outcome form includes the DR-learner [van der Laan and Luedtke, 2015, Kennedy, 2023b, Van Der Laan et al., 2023] and R-learner [Nie and Wager, 2021, Morzywolek et al., 2023] for the CATE $\theta_0(x) := \mu_0(1, x) - \mu_0(0, x)$. The DR-learner takes $w_{\eta_0}(z) = 1$ and

$$\zeta_{\eta_0}(z) = \mu_0(1, x) - \mu_0(0, x) + \frac{a - \pi_0(x)}{\pi_0(x)\{1 - \pi_0(x)\}}\{y - \mu_0(a, x)\},$$

where $\eta_0 = (\pi_0, \mu_0)$, $\mu_0(a, x) := E_0[Y \mid A = a, X = x]$, and $\pi_0(x) := P_0(A = 1 \mid X = x)$. The R-learner takes $w_{\eta_0}(z) = \{a - \pi_0(x)\}^2$ and $\zeta_{\eta_0}(z) = \{y - m_0(x)\}/\{a - \pi_0(x)\}$, where $\eta_0 = (\pi_0, m_0)$ and $m_0(x) := E_0[Y \mid X = x]$.

The next example shows that our framework also applies to smooth functionals of solutions to conditional moment restrictions, including potentially ill-posed inverse problems [Carrasco et al., 2007]. This class includes targets arising in nonparametric instrumental variables (NPIV) [Syrgkanis et al., 2019, Bennett et al., 2023a, 2025, van der Laan et al., 2025d], proximal causal inference [Zhang et al., 2023, Sverdrup and Cui, 2023, Cui et al., 2024], and offline reinforcement learning [Kallus and Uehara, 2020, 2022, van der Laan et al., 2025c,a].

Example 2 (Functionals in ill-posed inverse problems)**.** Consider the inverse-problem setting of Bennett et al. [2023b, 2025], Olivas-Martinez and Rotnitzky [2025], Ghassami et al. [2025]. Let $V \sim P_0$, let $X = X(V)$ and $Z = Z(V)$, and suppose h_0 satisfies

$$E_0\{g_1(V)h_0(X) - g_0(V) \mid Z\} = 0,$$

where g_1 and g_0 are known. With

$$(\mathcal{T}_0 h)(z) := E_0\{g_1(V)h(X) \mid Z = z\}, \qquad r_0(z) := E_0\{g_0(V) \mid Z = z\},$$

the restriction is $\mathcal{T}_0 h = r_0$. If $\mathcal{T}_0$ has a nontrivial null space, h_0 is identified only up to the equivalence class $[h_0] = \{h : h - h_0 \in \ker(\mathcal{T}_0)\}$. Thus the estimand is $\theta_0 = [h_0] \in \mathcal{H}$, where $\mathcal{H} := L^2(P_{0,X})/\ker(\mathcal{T}_0)$ is equipped with

$$\langle [h_1], [h_2] \rangle_{\mathcal{H}} := E_0\{(\mathcal{T}_0 h_1)(Z)(\mathcal{T}_0 h_2)(Z)\}.$$

This inner product is well defined and positive definite on $\mathcal{H}$. Accordingly, an identifiable functional must be invariant within equivalence classes; for a linear functional, this is equivalent to vanishing on $\ker(\mathcal{T}_0)$ [Babii, 2017, Babii and Florens, 2020, Bennett et al., 2025].

Let $\mathcal{N}$ be the Banach space of bounded linear operators from $L^2(P_{0,X})$ to $L^2(P_{0,Z})$, equipped with the operator norm, and view $\mathcal{T}_0 \in \mathcal{N}$ as a nuisance parameter. For a rep-

resentative h of $\theta = [h]$, Ghassami et al. [2025] derive the Neyman-orthogonal quadratic loss

$$\ell_{\eta_0}(V,h) := \frac{1}{2}\{(\mathcal{T}_0 h)(Z) - g_0(V)\}^2 + \{(\mathcal{T}_0 h)(Z) - r_0(Z)\}\{g_1(V)h(X) - (\mathcal{T}_0 h)(Z)\},$$

where $\eta_0 = (\mathcal{T}_0, r_0)$. Its risk equals

$$L_0(h,\eta_0) = \frac{1}{2}E_0[\{(\mathcal{T}_0 h)(Z) - r_0(Z)\}^2] + \frac{1}{2}E_0[\{g_0(V) - r_0(Z)\}^2],$$

so the induced objective on $\mathcal{H}$ is minimized at $\theta_0 = [h_0]$. Moreover, for directions u_1, u_2,

$$\partial_h^2 L_0(h,\eta_0)(u_1,u_2) = E_0\{(\mathcal{T}_0 u_1)(Z)(\mathcal{T}_0 u_2)(Z)\} = \langle [u_1], [u_2]\rangle_{\mathcal{H}}.$$

Thus the Hessian is the quotient-space inner product. For any smooth identifiable functional ψ_0, the Hessian Riesz representer is therefore

$$\alpha_0 \in \operatorname*{argmin}_{\alpha\in\mathcal{H}} \left\{ E_0[(\mathcal{T}_0\alpha)(Z)^2] - 2\,\dot{\psi}_0(\theta_0)(\alpha) \right\},$$

which exists only when $\dot{\psi}_0(\theta_0)$ is well defined on the quotient space, equivalently when it vanishes on $\ker(\mathcal{T}_0)$.

This template includes NPIV: with $V = (Y, W, Z_{\text{iv}})$, take $g_1 \equiv 1$, $g_0(V) = Y$, $X = W$, and $Z = Z_{\text{iv}}$, yielding $E_0\{h_0(W) - Y \mid Z_{\text{iv}}\} = 0$. It also includes regression-based bridge functions in proximal causal inference: with $V = (Y, A, W, Z_{\text{prox}}, X_0)$, take $g_1(V) = \mathbb{1}(A = a)$, $g_0(V) = \mathbb{1}(A = a)Y$, $X = (W, X_0)$, and $Z = (Z_{\text{prox}}, A, X_0)$, yielding $E_0[\mathbb{1}(A = a)\{h_0(W, X_0) - Y\} \mid Z_{\text{prox}}, A, X_0] = 0$. Finally, it covers value and Q-function estimation in offline reinforcement learning: with $V = (S, A, S')$, let $\mathcal{T}_0 = I - \gamma K_0$, where $(K_0 q)(a, s) := E_0\{\sum_{a'\in\mathcal{A}} \pi(a' \mid S')q(a', S') \mid A = a, S = s\}$. □

The final example illustrates that sequential regressions, as arise in longitudinal causal inference, optimal regime estimation, and nested structural models, can be treated as a single vector-valued M-estimand defined by a nuisance-dependent orthogonal loss [van der Laan and Robins, 2003, Robins, 2004, van der Laan and Luedtke, 2015, Luedtke et al., 2017, Chernozhukov et al., 2022b, Shirakawa et al., 2024].

Example 3 (Sequential regression as a coupled M-estimand)**.** Let $Z = (W_0, A_0, W_1, A_1, Y)$, let $\pi = (\pi_0, \pi_1)$ be a fixed intervention regime, let $p_t(a_t \mid w_t) = P_0(A_t = a_t \mid W_t = w_t)$ denote the observed propensity scores, and write $X_t = (W_t, A_t)$. For any f, write

$$\bar{f}_\pi(W_1) := \sum_{a_1} \pi_1(a_1 \mid W_1) f(W_1, a_1).$$

The target is the mean counterfactual outcome

$$\psi_0 = E_0\left[\sum_{a_0} \pi_0(a_0 \mid W_0)\theta_{0,1}(W_0, a_0)\right], \quad \theta_{0,1}(X_0) = E_0\{\bar{\theta}_{0,2,\pi}(W_1) \mid X_0\}, \quad \theta_{0,2}(X_1) = E_0(Y \mid X_1).$$

Let $\mathcal{T}_0 g(X_0) := E_0\{g(W_1) \mid X_0\}$. Then $\theta_{0,1} = \mathcal{T}_0\bar{\theta}_{0,2,\pi}$, and $\theta_0 = (\theta_{0,1}, \theta_{0,2})$ minimizes, over $\mathcal{H} = L^2(P_{0,X_0}) \times L^2(P_{0,X_1})$,

$$L_0(\theta) = \frac{1}{2}E_0\{Y - \theta_2(X_1)\}^2 + \frac{1}{2}E_0[\mathcal{T}_0\bar{\theta}_{2,\pi}(X_0) - \theta_1(X_0)]^2.$$

Let $\mathcal{N}$ be a Banach space of conditional-mean operators η from functions of W_1 to functions

of X_0, with $\eta_0 g(X_0) = E_0\{g(W_1) \mid X_0\}$. Define

$$\ell_\eta(Z,\theta) = \frac{1}{2}\{Y - \theta_2(X_1)\}^2 + \frac{1}{2}[\eta\bar{\theta}_{2,\pi}(X_0) - \theta_1(X_0)]^2 + [\eta\bar{\theta}_{2,\pi}(X_0) - \theta_1(X_0)]\{\bar{\theta}_{2,\pi}(W_1) - \eta\bar{\theta}_{2,\pi}(X_0)\}.$$

Then $E_0\ell_{\eta_0}(Z,\theta)$ is minimized at θ_0, and ℓ_η is Neyman orthogonal at (θ_0,η_0). The Hessian inner product is

$$\partial_\theta^2 L_0(\theta_0,\eta_0)(h,g) = E_0\{h_2(X_1)g_2(X_1)\} + E_0[\{(\mathcal{T}_0 h_2)(X_0) - h_1(X_0)\}\{(\mathcal{T}_0 g_2)(X_0) - g_1(X_0)\}]\,.$$

The Hessian Riesz representer $\alpha_0 = (\alpha_{0,1}, \alpha_{0,2})$ for ψ_0 is characterized by

$$\partial_\theta^2 L_0(\theta_0,\eta_0)(\alpha_0,h) = E_0\{\beta_0(X_0)h_1(X_0)\}, \qquad \beta_0(X_0) = \frac{\pi_0(A_0 \mid W_0)}{p_0(A_0 \mid W_0)}.$$

Equivalently, with $\rho_0(X_1) := \pi_1(A_1 \mid W_1)/p_1(A_1 \mid W_1)$,

$$\alpha_{0,2}(X_1) = \rho_0(X_1)E_0\{\beta_0(X_0) \mid W_1\}, \qquad \alpha_{0,1}(X_0) = \beta_0(X_0) + \eta_0\bar{\alpha}_{0,2,\pi}(X_0).$$

Therefore

$$-\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0) = \beta_0(X_0)\{\bar{\theta}_{0,2,\pi}(W_1) - \theta_{0,1}(X_0)\} + \alpha_{0,2}(X_1)\{Y - \theta_{0,2}(X_1)\},$$

which is the usual two-stage sequential regression correction [van der Laan and Gruber, 2012, Chernozhukov et al., 2022b]. □

4.3 Functional bias expansion and statistical efficiency

4.3.1 Differentiability conditions for loss functions and target parameters

This section introduces the smoothness conditions imposed on the risk functional $(\theta,\eta) \mapsto L_0(\theta,\eta)$ and the target functional ψ_0. We formulate these conditions in terms of Fréchet differentiability. At a high level, they require first- and second-order functional smoothness, together with Lipschitz continuity of the relevant derivatives, so that the remainder terms in the corresponding functional Taylor expansions can be controlled. The next subsection uses these conditions to derive a functional von Mises expansion for ψ_0 and, in turn, the efficient influence function of Ψ.

Functional Taylor expansions often require remainders to be small uniformly over the unit ball of $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$. When $\|\cdot\|_{\mathcal{H}}$ is an L^2-type norm, such uniform control may fail in infinite-dimensional settings. We therefore impose the remainder bounds uniformly over directions controlled by a stronger auxiliary norm $\rho_{\mathcal{H}}$, such as a supremum norm. In the conditions below, $\rho_{\mathcal{H}} : \mathcal{H} \to [0,\infty]$ denotes this auxiliary norm. For example, $\rho_{\mathcal{H}}$ may be an essential supremum norm, while $\|\cdot\|_{\mathcal{H}}$ may be an $L^2(P_0)$ norm or the Hessian-induced norm. All norms may depend on P_0.

A1) *(Well-defined population M-estimand).* For all $P \in \mathcal{P}$, the population risk admits a unique minimizer

$$\theta_P := \arg\min_{\theta\in\mathcal{H}} L_P(\theta,\eta_P).$$

Hereafter, let $\mathcal{H}_{\mathcal{P}} := \mathrm{conv}\{\theta_P : P \in \mathcal{P}\} \subset \mathcal{H}$ and $\mathcal{N}_{\mathcal{P}} := \mathrm{conv}\{\eta_P : P \in \mathcal{P}\} \subset \mathcal{N}$, where conv denotes the convex hull. Some conditions below are required only for $\theta \in \mathcal{H}_{\mathcal{P}}$ and $\eta \in \mathcal{N}_{\mathcal{P}}$.

Our next conditions concern the functional differentiability of the target and risk. A map $T : \mathcal{F} \to \mathcal{G}$ between normed spaces $(\mathcal{F}, \|\cdot\|_{\mathcal{F}})$ and $(\mathcal{G}, \|\cdot\|_{\mathcal{G}})$ is *Fréchet differentiable*

at $f_0 \in \mathcal{F}$ if there exists a continuous linear operator $\partial_f T(f_0) : \mathcal{F} \to \mathcal{G}$ such that $\|T(f_0 + h) - T(f_0) - \partial_f T(f_0)(h)\|_{\mathcal{G}} = o(\|h\|_{\mathcal{F}})$ as $\|h\|_{\mathcal{F}} \to 0$ [Rudin, 1991, van der Vaart and Wellner, 1996]. When $\mathcal{G} = \mathbb{R}$, the derivative $\partial_f T(f_0)$ lies in the *dual space* $\mathcal{F}^*$, the normed linear space of continuous linear functionals $B : \mathcal{F} \to \mathbb{R}$, equipped with the operator norm $\|B\|_{\mathcal{F}} := \sup_{f \in \mathcal{F}} |B(f)|$. Higher-order derivatives are defined recursively: T is *twice Fréchet differentiable* at f_0 if $f \mapsto \partial_f T(f)$ is Fréchet differentiable into $\mathcal{F}^*$, with $\partial_f^2 T(f_0) : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ given by this derivative. Since $\mathcal{N}$ need not be linear, Fréchet differentiability in η is understood after locally extending the relevant map to the ambient space $\widetilde{\mathcal{N}}$.

Our next set of assumptions ensures that $\psi_0(\theta)$ and the risk $L_0(\theta, \eta)$ admit sufficiently smooth local expansions in (θ, η).

A2) *(Hölder Smoothness of target functional):* The functional $\psi_0 : (\mathcal{H}, \|\cdot\|_{\mathcal{H}}) \to \mathbb{R}$ is Fréchet differentiable over $\mathcal{H}_{\mathcal{P}} \subset \mathcal{H}$ with Lipschitz continuous derivative $\theta \mapsto \dot{\psi}_0(\theta)$.

A3) *(Hölder smoothness of risk functional):*

i) *(First-order Fréchet differentiability)* For all $\theta \in \mathcal{H}_{\mathcal{P}}$ and $\eta \in \mathcal{N}_{\mathcal{P}}$, there exists a continuous linear operator $\dot{\ell}_\eta(\theta) : (\mathcal{H}, \|\cdot\|_{\mathcal{H}}) \to L^2(P_0)$ such that, for all $P \in \mathcal{P}$, the map $L_P(\cdot, \eta) : (\mathcal{H}, \|\cdot\|_{\mathcal{H}}) \to \mathbb{R}$ is Fréchet differentiable at θ with derivative satisfying $\partial_\theta L_P(\theta, \eta)(h) = P\, \dot{\ell}_\eta(\theta)(h)$ for all $h \in \mathcal{H}$.

ii) *(Second-order Fréchet differentiability)* For each $\eta \in \mathcal{N}_{\mathcal{P}}$, the map $\partial_\theta L_0(\cdot, \eta) : (\mathcal{H}, \|\cdot\|_{\mathcal{H}}) \to (\mathcal{H}, \rho_{\mathcal{H}})^*$ is Fréchet differentiable over $\mathcal{H}_{\mathcal{P}}$, where $\partial_\theta^2 L_0(\theta, \eta)(\cdot, \cdot)$ is symmetric.

iii) *(Lipschitz continuity of second derivative)* There exists a constant $C < \infty$ such that, for all $\theta, \theta' \in \mathcal{H}_{\mathcal{P}}$, $\eta, \eta' \in \mathcal{N}_{\mathcal{P}}$, and $h_1, h_2 \in \mathcal{H}$ with $\|h_1\|_{\mathcal{H}} + \rho_{\mathcal{H}}(h_2) \le 1$,

$$\left|\partial_\theta^2 L_0(\theta', \eta')(h_1, h_2) - \partial_\theta^2 L_0(\theta, \eta)(h_1, h_2)\right| \le C\left(\|\theta' - \theta\|_{\mathcal{H}} + \|\eta' - \eta\|_{\mathcal{N}}\right).$$

A4) *(Hölder Smoothness in nuisance parameter):*

i) The map $\eta \mapsto \partial_\theta L_0(\theta_0, \eta)$, defined on $\mathcal{N}_{\mathcal{P}}$, is Fréchet differentiable in η, with derivative $\partial_\eta \partial_\theta L_0(\theta_0, \eta) : \widetilde{\mathcal{N}} \to (\mathcal{H}, \rho_{\mathcal{H}})^*$.

ii) *(Lipschitz continuity of the cross-derivative).* There exists a constant $C < \infty$ such that, for all $\eta \in \mathcal{N}_{\mathcal{P}}$, $h \in \mathcal{H}$, and $g \in \widetilde{\mathcal{N}}$ satisfying $\eta + g \in \mathcal{N}_{\mathcal{P}}$ and $\|g\|_{\mathcal{N}} + \rho_{\mathcal{H}}(h) \le 1$,

$$|\partial_\eta \partial_\theta L_0(\theta_0, \eta)(g, h) - \partial_\eta \partial_\theta L_0(\theta_0, \eta_0)(g, h)| \le C\|\eta - \eta_0\|_{\mathcal{N}}.$$

We briefly explain the role of these conditions. Condition A1 ensures that the population M-estimand is well defined. For instance, if $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ is a Hilbert space, this follows when $\theta \mapsto L_P(\theta, \eta_P)$ is strictly convex, coercive, and lower semicontinuous [Ekeland and Temam, 1999]; strong convexity is a convenient sufficient condition. Condition A2 ensures that the target is locally linear in the M-estimand. Lipschitz continuity of $\theta \mapsto \dot{\psi}_0(\theta)$ yields, uniformly over local perturbations of θ_0,

$$\psi_0(\theta) = \psi_0(\theta_0) + \dot{\psi}_0(\theta_0)(\theta - \theta_0) + O\big(\|\theta - \theta_0\|_{\mathcal{H}}^2\big).$$

If ψ_0 is a continuous linear functional, then $\dot{\psi}_0(\theta)(h) = \psi_0(h)$. More generally, in many examples, $\dot{\psi}_0(\theta)(h) = E_0\{\dot{m}_\theta(Z,h)\}$, where $\dot{m}_\theta$ is the Gâteaux derivative of $m(z,\theta)$. Conditions A3i–A4 provide a local expansion of the risk gradient. Condition A3i, together with the definition of θ_0, gives

$$\partial_\theta L_0(\theta_0,\eta_0)(h) = 0 \qquad \text{for all } h \in \mathcal{H}.$$

The remaining conditions imply that, uniformly over h,

$$\begin{aligned}\partial_\theta L_0(\theta,\eta)(h) = {} & \partial_\theta^2 L_0(\theta_0,\eta_0)(\theta-\theta_0,h) + \partial_\eta\partial_\theta L_0(\theta_0,\eta_0)(\eta-\eta_0,h) \\ & + O\left[\rho_{\mathcal{H}}(h)\left\{\|\theta-\theta_0\|_{\mathcal{H}}^2 + \|\eta-\eta_0\|_{\mathcal{N}}^2\right\}\right].\end{aligned} \tag{4.5}$$

Thus, to first order, the plug-in bias is governed by the risk Hessian and by the effect of nuisance estimation on the risk gradient. If the loss is pointwise twice differentiable, then often $\partial_\theta^2 L_0(\theta,\eta)(h_1,h_2) = P_0\ddot{\ell}_\eta(\theta)(h_1,h_2)$. Because these assumptions are imposed on the risk rather than on the pointwise loss, they also cover nonsmooth losses, such as absolute-error and quantile losses, whenever the induced risk is sufficiently smooth.

To accommodate data-driven nuisance estimation, we use losses whose risk gradients are Neyman-orthogonal to nuisance perturbations [Robins et al., 1994, van der Laan and Robins, 2003, Chernozhukov et al., 2018a, Foster and Syrgkanis, 2023]. In view of (4.5), this condition sets the first-order nuisance contribution to the risk-gradient expansion to zero, so that nuisance estimation affects the estimating equation only through higher-order remainders. This permits flexible estimation of η_0, provided these remainders are controlled. We impose the following condition.

A5) *(Neyman orthogonality in the nuisance parameter):* For each $\eta \in \mathcal{N}_{\mathcal{P}}$ and $h \in \mathcal{H}$, $\partial_\eta\partial_\theta L_0(\theta_0,\eta_0)(\eta-\eta_0,h) = 0$.

Condition A5 is imposed only at θ_0, although many losses satisfy the stronger global condition $\partial_\eta\partial_\theta L_0(\theta,\eta_0)(\eta-\eta_0,h) = 0$ for all $\theta \in \mathcal{H}$, $\eta \in \mathcal{N}_{\mathcal{P}}$, and $h \in \mathcal{H}$ [Whitehouse et al., 2024]. Given a nonorthogonal loss, orthogonality at θ_0 can often be obtained by adding a correction term that is mean zero at η_0. This preserves the target risk at the truth, while changing the local nuisance sensitivity of the risk-gradient representation [Foster and Syrgkanis, 2023, Appendix D]. To see this, suppose

$$\partial_\eta\partial_\theta L_0(\theta_0,\eta_0)(\eta-\eta_0,h) = E_0\{a_{0,h}(Z)^\top(\eta-\eta_0)(Z)\}.$$

One then seeks, when possible, a corrected loss $\widetilde{\ell}$ whose population gradient at θ_0 satisfies

$$\partial_\theta E_0\{\widetilde{\ell}_{\theta_0,\eta}(Z)\}(h) = \partial_\theta L_0(\theta_0,\eta)(h) - E_0\{a_{0,h}(Z)^\top(\eta-\eta_0)(Z)\}.$$

Differentiating in η at η_0 gives $\partial_\eta\partial_\theta E_0\{\widetilde{\ell}_{\theta_0,\eta}(Z)\}(\eta-\eta_0,h) = 0$. Thus, the corrected loss defines the same target at the truth, while its estimating equation is Neyman-orthogonal. In many examples, the correction is obtained from the efficient influence function of the risk parameter $P \mapsto L_P(\theta,\eta_P)$, or of its gradient $P \mapsto \partial_\theta L_P(\theta_0,\eta_P)(h)$, when these maps are pathwise differentiable [Rubin and van der Laan, 2007, Yang et al., 2023, van der Laan et al., 2024a, Chen et al., 2026]. The corresponding empirical risk is often the associated one-step estimator [Bickel et al., 1993]. Profile losses provide another route to orthogonality [Murphy and Van der Vaart, 2000].

The next lemma gives simple sufficient conditions for Conditions A3 and A4. Let $\mathcal{H} = (L^\infty(\lambda))^{d_1}$ and $\widetilde{\mathcal{N}} = (L^\infty(\lambda))^{d_2}$, where λ is a measure on $\mathcal{Z} \subset \mathbb{R}^d$ dominating each $P \in \mathcal{P}$.

Equip these spaces with the norms

$$\|\theta\|_{\mathcal{H}} := \left\{ \int \|\theta(z)\|_{\mathbb{R}^{d_1}}^2 \, dP_0(z) \right\}^{1/2}, \qquad \|\eta\|_{\mathcal{N}} := \left\{ \int \|\eta(z)\|_{\mathbb{R}^{d_2}}^2 \, dP_0(z) \right\}^{1/2},$$

and write $\rho_{\mathcal{H}}(\theta) := \operatorname{ess\,sup}_{z \in \mathcal{Z}} \|\theta(z)\|_\infty$.

Lemma 4.7 (Pointwise smooth losses)**.** *Suppose $\ell_\eta(\theta, z) = l\{\theta(z), \eta(z), z\}$, where $l : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^d \to \mathbb{R}$. Let $\mathcal{C}$ be the closure of the set of all (a, b, z) such that $z \in \mathcal{Z}$, $a = \theta(z)$, $b = \eta(z)$, $\theta \in \mathcal{H}_\mathcal{P}$, and $\eta \in \mathcal{N}_\mathcal{P}$. Assume that $\mathcal{C}$ is compact and that, on $\mathcal{C}$, l is twice continuously differentiable in a and once continuously differentiable in b. Assume further that $\partial_a^2 l$ and $\partial_a \partial_b l$ are Lipschitz in (a, b), uniformly in z, and that l, $\partial_a l$, $\partial_a^2 l$, and $\partial_a \partial_b l$ are jointly continuous in (a, b, z). Then Conditions A3 and A4 hold. The corresponding derivatives are $\dot{\ell}_\eta(\theta, z)(h) = \partial_a l\{\theta(z), \eta(z), z\}^\top h(z)$,*

$$\partial_\theta^2 L_0(\theta, \eta)(h_1, h_2) = P_0\Big[h_1(Z)^\top \partial_a^2 l\{\theta(Z), \eta(Z), Z\} h_2(Z)\Big],$$

$$\partial_\eta \partial_\theta L_0(\theta, \eta)(g, h) = P_0\Big[h(Z)^\top \partial_a \partial_b l\{\theta(Z), \eta(Z), Z\} g(Z)\Big].$$

In Appendix 4.D, we apply Lemma 4.7 to the orthogonal logistic loss from Example 1b. We first illustrate the conditions in the running example.

Example 1a (continued)**.** Consider the Riesz loss $\ell(\theta, z) := \theta(x)^2/2 - r(z, \theta)$, where $r : \mathcal{Z} \times \mathcal{H} \to \mathbb{R}$ is linear in its second argument. Assume each $P \in \mathcal{P}$ is dominated by P_0 and that, for some $0 < c < C < \infty$, $c\|f\|_{L^2(P_0)} \le \|f\|_{L^2(P)} \le C\|f\|_{L^2(P_0)}$. Let $\mathcal{H} = L^2(P_0)$, with $\|\cdot\|_{\mathcal{H}}$ the $L^2(P_0)$ norm and $\rho_{\mathcal{H}}$ the P_0-essential supremum norm. Then Condition A1 holds because $\theta \mapsto E_P\{\theta(X)^2/2 - r(Z, \theta)\}$ is strongly convex and continuous on $\mathcal{H}$.

Condition A2 holds with $\dot{\psi}_0 = \psi_0$ if $\theta \mapsto m(Z, \theta)$ is almost surely linear and $\sup_{\theta \in \mathcal{H}} |E_0\{m(Z, \theta)\}|/\|\theta\|_{\mathcal{H}} < \infty$. It also holds when $m(Z, \theta) = g\{\theta(Z)\}$ for a differentiable $g : \mathbb{R} \to \mathbb{R}$ with Lipschitz derivative $\dot{g}$; in this case $\dot{\psi}_0(\theta)(h) = E_0[\dot{g}\{\theta(Z)\}h(Z)]$ [Qiu et al., 2021, Section 4.1]. The loss satisfies Condition A3i with $\dot{\ell}(\theta)(h, z) = h(x)\theta(x) - r(z, h)$, and Condition A3ii with $\partial_\theta^2 L_0(\theta, \eta)(h_1, h_2) = \langle h_1, h_2 \rangle_{\mathcal{H}}$. Conditions A4 and A5 hold trivially, since the loss has no nuisance component. □

4.3.2 Functional von Mises expansion

This section develops a functional von Mises expansion for $\psi_0(\theta_0)$, which provides the basis for the asymptotic analysis of the autoDML estimators [Von Mises, 1947, Bickel et al., 1993].

Let $\overline{\mathcal{H}}$ denote the completion of $\mathcal{H}$ under $\|\cdot\|_{\mathcal{H}}$. The expansion involves the Hessian Riesz representer of the linearization of ψ_0. The following condition ensures that this representer exists.

A6) *(Hessian norm equivalence).* There exist constants $0 < \kappa_1 \le \kappa_2 < \infty$ such that

$$\kappa_1 \|h\|_{\mathcal{H}}^2 \le \partial_\theta^2 L_0(\theta_0, \eta_0)(h, h) \le \kappa_2 \|h\|_{\mathcal{H}}^2 \qquad \text{for all } h \in \mathcal{H}.$$

Condition A6 makes the Hessian norm equivalent to $\|\cdot\|_{\mathcal{H}}$. Together with Condition A2, this implies that $\dot{\psi}_0(\theta_0)$ is continuous in the Hessian norm. Hence, by the Riesz representation theorem, there exists a unique $\alpha_0 \in \overline{\mathcal{H}}$ satisfying (4.2). The condition is automatic if $\|\cdot\|_{\mathcal{H}}$

is chosen as the Hessian-induced norm $\{\partial_\theta^2 L_0(\theta_0,\eta_0)(h,h)\}^{1/2}$, provided the Hessian inner product is positive definite. If the Hessian is only positive semidefinite, one may instead work on the quotient space that identifies directions with zero Hessian seminorm, as in Example 2.

We say that ψ_0 is linear at θ_0 if $\psi_0(\theta) = \psi_0(\theta_0) + \dot{\psi}_0(\theta_0)(\theta - \theta_0)$. We say that $L_0(\cdot,\eta_0)$ is quadratic at θ_0 if

$$L_0(\theta,\eta_0) = L_0(\theta_0,\eta_0) + \partial_\theta L_0(\theta_0,\eta_0)(\theta-\theta_0) + \frac{1}{2}\partial_\theta^2 L_0(\theta_0,\eta_0)(\theta-\theta_0,\theta-\theta_0).$$

Theorem 4.9 (Functional von Mises expansion)**.** *Suppose Conditions A1–A6 hold. Then, for each $\eta \in \mathcal{N}_\mathcal{P}$, $\theta \in \mathcal{H}_\mathcal{P}$, and $\alpha \in \mathcal{H}$ with $\rho_\mathcal{H}(\alpha) < M$,*

$$\begin{aligned}\psi_0(\theta) - \psi_0(\theta_0) = P_0\dot{\ell}_\eta(\theta)(\alpha) + \partial_\theta^2 L_0(\theta_0,\eta_0)(\alpha_0-\alpha,\theta-\theta_0) \\ + O\big(\|\theta-\theta_0\|_\mathcal{H}^2\big) + O(\|\eta-\eta_0\|_\mathcal{N}\|\theta-\theta_0\|_\mathcal{H}) + O\big(\|\eta-\eta_0\|_\mathcal{N}^2\big)\,,\end{aligned}$$

where the constants in the $O(\cdot)$ terms depend only on M and the constants in the preceding conditions. If ψ_0 is linear at θ_0 and $L_0(\cdot,\eta_0)$ is quadratic at θ_0, then the $O(\|\theta-\theta_0\|_\mathcal{H}^2)$ term vanishes.

Theorem 4.9 shows that the leading plug-in error $\psi_0(\theta_n) - \psi_0(\theta_0)$ is $P_0\dot{\ell}_{\eta_0}(\theta_n)(\alpha_0)$, up to second-order terms. With estimators η_n of η_0 and α_n of α_0, this term is approximated by $P_0\dot{\ell}_{\eta_n}(\theta_n)(\alpha_n)$, again up to second-order terms. This suggests debiasing the plug-in estimator $\psi_n(\theta_n)$ either explicitly, by adding the empirical correction $-P_n\dot{\ell}_{\eta_n}(\theta_n)(\alpha_n)$, or implicitly, by constructing θ_n so that $P_n\dot{\ell}_{\eta_n}(\theta_n)(\alpha_n) = 0$. The next section uses this expansion to construct autoDML estimators for $\Psi(P_0)$.

If the target is linear and the risk is quadratic, Theorem 4.9 contains only the mixed-bias remainders $\partial_\theta^2 L_0(\theta_0,\eta_0)(\alpha_0-\alpha,\theta-\theta_0)$ and $\|\eta-\eta_0\|_\mathcal{N}\|\theta-\theta_0\|_\mathcal{H}$. The second term can be removed under a stronger, global form of Neyman orthogonality: namely, if $\partial_\eta\partial_\theta L_0(\theta,\eta_0)(\eta-\eta_0,h) = 0$ for all $\theta \in \mathcal{H}$, $\eta \in \mathcal{N}$, and $h \in \mathcal{H}$, together with a suitable Lipschitz condition; see Theorem C.9 in Appendix C.3. This extends the familiar double robustness of von Mises expansions for linear functionals of regression functions under squared-error loss to linear functionals of general pseudo-outcome regressions (Example 1c). In this case, the estimand also admits the dual representation $\Psi(P_0) = -\partial_\theta L_0(0,\eta_0)(\alpha_0) = -E_0\{\dot{\ell}_{\eta_0}(0)(\alpha_0)(Z)\}$.

Example 1a (continued)**.** Condition A6 holds since $\partial_\theta^2 L_0(\theta,\eta)(h,h) = \|h\|_\mathcal{H}^2$. For a linear functional ψ_0 of the Riesz representer θ_0, the term $-\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0)$ equals $z \mapsto r(z,\alpha_0) - \alpha_0(x)\theta_0(x)$, where $\alpha_0 := \operatorname{argmin}_{\alpha\in\mathcal{H}} E_0[\{\alpha(X)\}^2 - 2\,m(Z,\alpha)]$. When θ_0 is the outcome regression, we have $-\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0)(z) = \alpha_0(x)\{y-\theta_0(x)\}$. The von Mises expansion is doubly robust, satisfying $\psi_0(\theta) - \psi_0(\theta_0) - P_0\dot{\ell}_\eta(\theta)(\alpha) = \langle\theta-\theta_0,\alpha_0-\alpha\rangle_{L^2(P_0)}$. □

4.3.3 Pathwise differentiability and efficiency

The von Mises expansion above identifies the first-order correction needed to debias plug-in estimators of $\psi_0(\theta_0)$. We next connect this linearization to semiparametric efficiency by establishing pathwise differentiability of Ψ and deriving its efficient influence function [Bickel et al., 1993].

The role of the Hessian Riesz representer α_0 can be understood from a local perturbation argument. Consider a path $\theta_\varepsilon = \theta_0 + \varepsilon h$, with $h \in \overline{\mathcal{H}}$. Along this path,

$$\frac{|\dot{\psi}_0(\theta_0)(h)|^2}{\partial_\theta^2 L_0(\theta_0, \eta_0)(h, h)}$$

measures the squared first-order change in the target, normalized by the local curvature of the loss. Since

$$\dot{\psi}_0(\theta_0)(h) = \partial_\theta^2 L_0(\theta_0, \eta_0)(\alpha_0, h),$$

Cauchy–Schwarz implies that this ratio is maximized when h is proportional to α_0. Thus, α_0 is the local direction to which the target is most sensitive relative to the loss geometry. The theorem below shows that the same representer enters the efficient influence function through the loss derivative, and hence determines the nonparametric efficiency bound.

A7) *(Functional continuity):* $P \mapsto \Psi(P)$, $P \mapsto \theta_P$ and $P \mapsto \eta_P$ are Lipschitz continuous maps at P_0 with domain $\mathcal{P}$ equipped with the Hellinger distance and the range equipped with $|\cdot|$, $\|\cdot\|_{\mathcal{H}}$ and $\|\cdot\|_{\mathcal{N}}$, respectively. In addition, $(\eta, \theta) \mapsto \dot{\ell}_\eta(\cdot, \theta)$ and $(\eta, \theta) \mapsto m(\cdot, \theta)$ are continuous maps with domain $\mathcal{N} \times \mathcal{H}$ equipped with the total norm $(\eta, \theta) \mapsto \|\eta\|_{\mathcal{N}} + \|\theta\|_{\mathcal{H}}$ and range $L^2(P_0)$.

A8) *(Boundedness):* $\sup_{P \in \mathcal{P}} \rho_{\mathcal{H}}(\alpha_P) < \infty$, where $\alpha_P \in \overline{\mathcal{H}}$ denotes the Hessian Riesz representer under P.

Condition A7 requires $P \mapsto \Psi(P)$, $P \mapsto \theta_P$, and $P \mapsto \eta_P$ to be Hellinger Lipschitz, together with $L^2(P_0)$ continuity of the loss and target gradients. Analogous Hellinger-continuity conditions have been verified for common regression-type and causal parameters [Luedtke and Chung, 2024, Luedtke, 2024, Chen et al., 2026]. Appendix 4.C gives a sufficient condition for the M-estimand map $P \mapsto \theta_P$. We emphasize, however, that the von Mises expansion in Theorem 4.9 can still yield asymptotically linear estimators even without pathwise differentiability, although such estimators may be irregular. This phenomenon can arise in ill-posed inverse problems [Smucler et al., 2025, Bennett et al., 2025, van der Laan et al., 2025d], where the relevant nuisance components need not be Hellinger smooth.

Theorem 4.11 (Pathwise differentiability)**.** *If Conditions A1–A8 hold, then $\Psi : \mathcal{P} \to \mathbb{R}$ is pathwise differentiable at P_0, with efficient influence function*

$$\chi_0(z) = -\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0)(z) + m(z, \theta_0) - \Psi(P_0).$$

The theorem shows that the Hessian Riesz representer α_0 determines the component of the nonparametric efficient influence function arising from estimation of the M-estimand. This mirrors the role of the canonical gradient in classical semiparametric theory, but in the space of M-estimands rather than in the tangent space of the statistical model. In a likelihood model, the canonical gradient χ_0 represents the target derivative over distributional perturbations and is the score of a least favorable density submodel [Stein, 1956, Bickel et al., 1993]. By contrast, α_0 represents the target derivative with respect to perturbations of the M-estimand, under the Hessian inner product induced by the risk. Thus, χ_0 and α_0 are analogous Riesz representers in different spaces: χ_0 for perturbations of the data-generating distribution, and α_0 for perturbations of the induced M-estimand.

When $\mathcal{H}$ is imposed as a structural restriction rather than treated as a working model, the influence function in Theorem 4.11 need not be efficient. The theorem gives the nonparametric EIF and efficiency bound for the loss-induced projection parameter Ψ on the nonparametric model $\mathcal{M}$, where $\mathcal{H}$ only defines the projection target. For a restricted submodel in which the true M-estimand is assumed to lie in $\mathcal{H}$, the same expression is generally only an influence function for the restricted target. It is efficient only if it lies in the restricted tangent space; otherwise, the restricted-model EIF is its projection onto that space.

To make this distinction precise, let $\theta_{P,\mathrm{np}}$ be the unconstrained M-estimand obtained by minimizing the population risk over a nonparametric ambient space, such as $L^2(P_0)$, and define

$$\mathcal{M}_{\mathcal{H}} := \{P \in \mathcal{M} : \theta_{P,\mathrm{np}} \in \mathcal{H}\}.$$

When $P_0 \in \mathcal{M}_{\mathcal{H}}$, two losses can agree on the $\mathcal{H}$-constrained minimizer, and hence on the target value, while inducing different loss-based extensions to the larger model $\mathcal{M}$. The corresponding nonparametric EIFs and efficiency bounds may therefore differ. This can occur, for example, when $\mathcal{H}$ encodes a correctly specified linear regression model: ordinary squared-error loss is restricted-model efficient under homoskedastic errors, whereas inverse conditional-variance weighted squared-error loss recovers the semiparametric efficiency bound under heteroskedasticity [Chamberlain, 1987, Bickel et al., 1993]. The log-likelihood loss is another special case, often inducing an extension whose nonparametric EIF coincides with the restricted-model EIF, and hence attains the sharp efficiency bound under $\mathcal{M}_{\mathcal{H}}$ [van der Laan et al., 2023, Section 3].

Theorem 4.11 also clarifies how to construct a targeted minimum loss-based estimator for a general M-estimand, as developed in Section 4.4.2. Classical TMLE updates an initial density estimate along an EIF-based least favorable submodel, typically using the log-likelihood loss [Laan and Rubin, 2006, van der Laan and Rose, 2011]. For M-estimation, the analogous update starts from an initial M-estimator and fluctuates it in the Hessian Riesz direction α_0 to solve the empirical EIF equation. Thus, the theorem identifies the Hessian Riesz representer as playing the role of the "clever covariate" in targeted minimum loss estimation [van der Laan and Rose, 2011].

4.4 Automatic debiased machine learning

4.4.1 General form of the estimator

The expansion in Theorem 4.9 suggests a common template for autoDML estimation: start from a plug-in estimator and debias it by adding the empirical mean of an estimated EIF correction. In the present notation, this leads to estimators of the form

$$\widehat{\psi}_n = \psi_n(\widehat{\theta}_n) - P_n \dot{\ell}_{\eta_n}(\widehat{\theta}_n)(\alpha_n), \qquad \psi_n(\widehat{\theta}_n) = \frac{1}{n}\sum_{i=1}^{n} m(Z_i, \widehat{\theta}_n),$$

where η_n, $\widehat{\theta}_n$, and α_n estimate η_0, θ_0, and α_0, respectively. Here $\widehat{\theta}_n$ denotes a generic estimator, which may be either an initial estimator or a targeted version of one. The correction can be applied explicitly, as in the one-step estimator, or made to vanish by construction, as in targeted minimum loss-based and sieve estimators, by choosing $\widehat{\theta}_n$ so that the empirical derivative in the estimated Riesz direction equals zero.

The one-step estimator is the special case with $\widehat{\theta}_n = \theta_n$, where θ_n is the initial estimator:

$$\widehat{\psi}_n^{\mathrm{dml}} = \psi_n(\theta_n) - P_n \dot{\ell}_{\eta_n}(\theta_n)(\alpha_n).$$

The subtracted term estimates the leading plug-in bias identified by Theorem 4.9.

The procedure is automatic in the sense that, once the target map and orthogonal loss are specified, inference reduces to estimating η_0, the M-estimand θ_0, and the Hessian Riesz representer α_0. To obtain strong theoretical guarantees under weak complexity conditions, we recommend cross-fitting, a standard device for mitigating overfitting and relaxing empirical-process requirements in nuisance estimation [van der Laan et al., 2011, Chernozhukov et al., 2018a]. Cross-fitted versions are given in Appendix 4.A.

4.4.2 Automatic targeted minimum loss-based estimation

A limitation of the one-step estimator $\widehat{\psi}_n^{\mathrm{dml}}$ is that it debiases the plug-in estimator by adding a scalar correction outside the model. Even when the initial plug-in estimator $\psi_n(\theta_n)$ respects the constraints encoded by $\mathcal{H}$, the corrected estimate $\widehat{\psi}_n^{\mathrm{dml}} = \psi_n(\theta_n) - P_n \dot{\ell}_{\eta_n}(\theta_n)(\alpha_n)$ need not be representable as $\psi_n(\theta)$ for any $\theta \in \mathcal{H}$. This matters when $\mathcal{H}$ encodes known structural constraints, whether through the loss, the parametrization, or the constraint set itself, such as probability constraints, monotonicity, or bounds implied by bounded outcomes.

Targeted minimum loss-based estimation (TMLE) addresses this by debiasing the nuisance estimator itself rather than adding an explicit correction to the final scalar estimate [Laan and Rubin, 2006, van der Laan and Rose, 2011]. Starting from θ_n, TMLE performs a loss-based update within $\mathcal{H}$ that solves the relevant empirical efficient estimating equation. In our setting, the automatic targeted minimum loss-based estimator (autoTMLE) is

$$\widehat{\psi}_n^{\mathrm{tmle}} := \psi_n(\theta_n^*) = P_n m(\cdot, \theta_n^*),$$

where θ_n^* is chosen to solve

$$P_n \dot{\ell}_{\eta_n}(\theta_n^*)(\alpha_n) = 0. \tag{4.6}$$

Thus, the one-step correction vanishes at $(\eta_n, \theta_n^*, \alpha_n)$, so the resulting targeted plug-in estimator retains the substitution property in finite samples while achieving first-order debiasing and efficiency.

There are several ways to construct such a targeted estimator. A simple construction treats α_n as fixed and uses a one-dimensional fluctuation through the initial estimator θ_n:

$$\theta_n^* = \theta_n + \varepsilon_n \alpha_n, \qquad \varepsilon_n = \operatorname*{argmin}_{\varepsilon \in \mathbb{R}} \sum_{i=1}^{n} \ell_{\eta_n}(Z_i, \theta_n + \varepsilon \alpha_n).$$

Since $\mathcal{H}$ is linear, this fluctuation remains in $\mathcal{H}$, and the first-order condition for ε_n is exactly (4.6). Thus, autoTMLE is empirical risk minimization along an estimated least favorable submodel for ψ_0: it moves θ_n in the estimated Hessian Riesz direction α_n, the direction of steepest target change per unit local curvature of the risk. If α_n depends on the current M-estimand estimate, one can iterate the targeting step, recomputing α_n after each update, until the corresponding score equation is solved [Laan and Rubin, 2006]. Universal least favorable submodels provide another implementation, updating the fluctuation direction along the path so that the estimated efficient score remains aligned with the current M-estimand estimate [van der Laan and Gruber, 2016].

At first order, autoTMLE and the one-step estimator use the same debiasing direction

but implement debiasing differently: the one-step estimator adds the linear correction directly, whereas autoTMLE chooses the targeting step by empirical risk minimization. To see this, suppose that $\ell_\eta(\theta)$ is twice differentiable in θ and that $m(z,\theta)$ is differentiable. The autoTMLE update chooses ε_n by empirical risk minimization along the fluctuation submodel $\theta_n + \varepsilon\alpha_n$. A quadratic expansion of $\ell_{\eta_n}(Z, \theta_n + \varepsilon\alpha_n)$ around $\varepsilon = 0$ gives the Newton approximation

$$\theta_n^* \approx \theta_n - \frac{P_n \dot{\ell}_{\eta_n}(\theta_n)(\alpha_n)}{P_n \ddot{\ell}_{\eta_n}(\theta_n)(\alpha_n, \alpha_n)} \alpha_n.$$

A first-order expansion of m then gives

$$\widehat{\psi}_n^{\text{tmle}} \approx P_n m(\cdot, \theta_n) - \frac{P_n \dot{m}_{\theta_n}(\cdot, \alpha_n)}{P_n \ddot{\ell}_{\eta_n}(\theta_n)(\alpha_n, \alpha_n)} P_n \dot{\ell}_{\eta_n}(\theta_n)(\alpha_n). \tag{4.7}$$

Thus, the local linear approximation to the autoTMLE update has the algebraic form of a one-step correction with a data-adaptive scalar rescaling: the loss-based targeting step rescales the explicit one-step correction by $P_n \dot{m}_{\theta_n}(\cdot, \alpha_n)/P_n \ddot{\ell}_{\eta_n}(\theta_n)(\alpha_n, \alpha_n)$. Equivalently, this is the one-step estimator that would be obtained using the scalar-calibrated representer $c\alpha_n$, where c minimizes the empirical Riesz risk $c \mapsto P_n \ddot{\ell}_{\eta_n}(\theta_n)(c\alpha_n, c\alpha_n)/2 - P_n \dot{m}_{\theta_n}(\cdot, c\alpha_n)$ over $\{c\alpha_n : c \in \mathbb{R}\}$. By the Hessian Riesz representer identity, the population analogue of this scaling factor equals one when $\alpha_n = \alpha_0$. Hence the linearized autoTMLE recovers the usual one-step correction asymptotically under consistent nuisance estimation. In finite samples, the scaling factor lets the empirical loss choose the size of the correction, which can stabilize the update. The approximation is exact when the risk is quadratic and the target is linear, as illustrated in the next example.

Example 1a (continued)**.** For a linear functional of a Riesz representer θ_0, the one-step autoDML estimator is

$$\widehat{\psi}_n^{\text{dml}} := P_n m(\cdot, \theta_n) + P_n\{r(Z, \alpha_n(X)) - \theta_n(X)\alpha_n(X)\},$$
$$\widehat{\psi}_n^{\text{tmle}} = P_n m(\cdot, \theta_n) + \delta_n^* P_n\{r(Z, \alpha_n(X)) - \theta_n(X)\alpha_n(X)\}, \qquad \delta_n^* := \frac{P_n m(\cdot, \alpha_n)}{P_n\{\alpha_n^2\}}.$$

In the special case where θ_0 is the outcome regression, the autoDML correction becomes $P_n \alpha_n(X)\{Y - \theta_n(X)\}$. The autoTMLE estimator is obtained by fluctuating θ_n along α_n, with $\varepsilon_n = \operatorname{argmin}_\varepsilon P_n[\{\theta_n(X) + \varepsilon\alpha_n(X)\}^2 - 2\varepsilon r(Z, \alpha_n(X))]$. Thus, autoTMLE rescales the one-step correction by the data-adaptive normalization factor δ_n^*. □

4.4.3 Plug-in sieve M-estimation and automatic undersmoothing

A second route to debiased plug-in estimation is to compute the M-estimator over a sequence of increasingly rich finite-dimensional sieve spaces. When the sieve is sufficiently rich, the empirical score equations can eliminate the leading correction, so the resulting estimator is a pure plug-in estimator rather than an explicitly corrected one-step estimator. This is the idea behind sieve M-estimators [Shen, 1997, Chen, 2007, Chen and Liao, 2014, Qiu et al., 2021, van der Laan and Rose, 2021, van der Laan et al., 2022]. Sieve methods are especially attractive for low-dimensional targets, such as a dose-response curve or a low-dimensional CATE function. They can be less practical in high-dimensional problems, where constructing a sieve with adequate approximation guarantees is difficult. Orthogonal losses make this approach more flexible by allowing the sieve to target the primary M-estimand, while other

nuisance components enter through the orthogonalized loss.

Let $\mathcal{H}_1 \subset \mathcal{H}_2 \subset \cdots \subset \mathcal{H}$ be a nested sequence of finite-dimensional subspaces whose union is dense in $\mathcal{H}$. For each k, define the sieve M-estimator

$$\theta_{n,k} := \operatorname*{argmin}_{\theta \in \mathcal{H}_k} P_n \ell_{\eta_n}(Z, \theta), \qquad \widehat{\psi}^{\text{sieve}}_{n,k} := P_n m(\cdot, \theta_{n,k}).$$

The autoSieve estimator is $\widehat{\psi}^{\text{sieve}}_n := \widehat{\psi}^{\text{sieve}}_{n,k(n)}$, where $k(n)$ is chosen as described below.

The sieve plug-in estimator is debiased because its empirical score equations hold in every direction in the sieve space, including directions that approximate the Hessian Riesz representer. Specifically,

$$P_n \dot{\ell}_{\eta_n}(\theta_{n,k})(\alpha) = 0 \qquad \text{for all } \alpha \in \mathcal{H}_k.$$

In particular, the score equation holds for the Hessian projection

$$\alpha_{0,k} := \operatorname*{argmin}_{\alpha \in \mathcal{H}_k} \partial^2_\theta L_0(\theta_0, \eta_0)(\alpha - \alpha_0, \alpha - \alpha_0).$$

Therefore, $\widehat{\psi}^{\text{sieve}}_{n,k}$ has the same form as the general debiased template with $(\eta_n, \theta_{n,k}, \alpha_{0,k})$, but with the correction term equal to zero. To remove first-order plug-in bias, the sieve dimension must be large enough to approximate both θ_0 and α_0. This typically requires undersmoothing, so that $k(n)$ grows faster than would be optimal for estimating θ_0 alone.

We select $k(n)$ by an automatic undersmoothing rule. Using an independent validation sample $\{Z'_i\}_{i=1}^{n'}$, first choose $k_\theta(n) := \operatorname{argmin}_k P'_{n'} \ell_{\eta_n}(Z, \theta_{n,k})$, which targets estimation of θ_0. Next, with $k_\theta = k_\theta(n)$, estimate the projected representer by

$$\alpha_{n,k} := \operatorname*{argmin}_{\alpha \in \mathcal{H}_k} P_n \left[\ddot{\ell}_{\eta_n}(\theta_{n,k_\theta})(\alpha, \alpha)(Z) - 2\, \dot{m}_{\theta_{n,k_\theta}}(Z, \alpha) \right],$$

and choose

$$k_\alpha(n) := \operatorname*{argmin}_k P'_{n'} \left[\ddot{\ell}_{\eta_n}(\theta_{n,k_\theta})(\alpha_{n,k}, \alpha_{n,k})(Z) - 2\, \dot{m}_{\theta_{n,k_\theta}}(Z, \alpha_{n,k}) \right].$$

We then set $k(n) := \max\{k_\theta(n), k_\alpha(n)\}$. The first dimension selects the sieve size needed to estimate the M-estimand, while the second selects the size needed to approximate the Hessian representer. Taking the maximum automatically undersmooths the plug-in estimator when α_0 is more complex than θ_0, a common source of first-order bias [Qiu et al., 2021]. The validation sample can be replaced by sample-splitting or cross-validation.

4.5 Theoretical results

4.5.1 Asymptotic theory for autoDML

In this section, we establish regularity, asymptotic linearity, and semiparametric efficiency for estimators constructed from the general autoDML template in Section 4.4.1. The result applies to any estimator of the form

$$\widehat{\psi}_n = \psi_n(\widehat{\theta}_n) - P_n \dot{\ell}_{\eta_n}(\widehat{\theta}_n)(\alpha_n),$$

with the convention that the correction term may vanish by construction. The one-step, autoTMLE, and autoSieve estimators are obtained by different choices of $\widehat{\theta}_n$ and α_n. Denote the estimated influence function corresponding to $(\widehat{\theta}_n, \eta_n, \alpha_n)$ by

$$\chi_n := m(\cdot, \widehat{\theta}_n) - P_n m(\cdot, \widehat{\theta}_n) - \dot{\ell}_{\eta_n}(\widehat{\theta}_n)(\alpha_n).$$

Our main theorem uses the following conditions.

B1) *Bounded estimators:* With probability tending to one, $\eta_n \in \mathcal{N}_{\mathcal{P}}$, $\widehat{\theta}_n \in \mathcal{H}_{\mathcal{P}}$, and $\rho_{\mathcal{H}}(\alpha_n) = O_p(1)$.

B2) *Nuisance estimation rate:* $\|\eta_n - \eta_0\|_{\mathcal{N}} = o_p(n^{-1/4})$.

B3) *M-estimation rate:* One of the following holds:

(i) ψ_0 is linear and $L_0(\cdot, \eta_0)$ is quadratic at θ_0;
(ii) $\|\widehat{\theta}_n - \theta_0\|_{\mathcal{H}} = o_p(n^{-1/4})$.

B4) *Doubly robust rate:* $\partial^2_\theta L_0(\theta_0, \eta_0)(\alpha_0 - \alpha_n, \widehat{\theta}_n - \theta_0) + \|\widehat{\theta}_n - \theta_0\|_{\mathcal{H}} \|\eta_n - \eta_0\|_{\mathcal{N}} = o_p(n^{-1/2})$.

B5) *Empirical process condition:* $(P_n - P_0)(\chi_n - \chi_0) = o_p(n^{-1/2})$.

We briefly discuss the conditions and give a fuller treatment in Appendix 4.C. Conditions B2 and B3 require either the linear-quadratic structure in Condition B3(i), or consistency of η_n and $\widehat{\theta}_n$ at rates faster than $n^{-1/4}$. Such rates are standard in semiparametric statistics and can be achieved under appropriate conditions by generalized additive models [Hastie and Tibshirani, 1987], the highly adaptive lasso [van der Laan, 2017], gradient-boosted trees [Schuler et al., 2023], and neural networks [Farrell et al., 2018]. Condition B4 is a mixed-bias condition; its first term is $o_p(n^{-1/2})$, for example, when $\|\alpha_n - \alpha_0\|_{\mathcal{H}} \|\widehat{\theta}_n - \theta_0\|_{\mathcal{H}} = o_p(n^{-1/2})$ under Condition A3ii. Condition B5 is a standard empirical-process condition, which holds when $\|\chi_n - \chi_0\|_{L^2(P_0)} = o_p(1)$ and the nuisance estimators are constructed using appropriate cross-fitting techniques [van der Laan et al., 2011, Chernozhukov et al., 2018a].

Theorem 4.13 (Asymptotic linearity and efficiency)**.** *Assume Conditions A1–A6 and B1–B5. Then*

$$\widehat{\psi}_n - \Psi(P_0) = P_n \chi_0 + o_p(n^{-1/2}),$$

and hence

$$n^{1/2}\{\widehat{\psi}_n - \Psi(P_0)\} \rightsquigarrow N\{0, \mathrm{var}_0(\chi_0(Z))\}.$$

If Conditions A7 and A8 also hold, then $\widehat{\psi}_n$ is regular and efficient for $\Psi(P_0)$.

The theorem applies directly to the one-step estimator by taking $\widehat{\theta}_n = \theta_n$. The autoTMLE and autoSieve estimators are obtained from the same template with targeted or sieve M-estimators for $\widehat{\theta}_n$, for which the empirical correction term vanishes by construction. Hence Theorem 4.13 gives the following corollary.

Corollary 4.14 (Asymptotic efficiency of TMLE and sieve estimators)**.** *Assume Conditions A1–A8.*

(i) If Conditions B1–B5 hold with $(\eta_n, \theta^*_n, \alpha_n)$ in place of $(\eta_n, \widehat{\theta}_n, \alpha_n)$, then $\widehat{\psi}^{\mathrm{tmle}}_n$ is regular, asymptotically linear, and efficient for $\Psi(P_0)$, with influence function χ_0.

(ii) If Conditions B1–B5 hold with $(\eta_n, \theta_{n,k(n)}, \alpha_{0,k(n)})$ *in place of* $(\eta_n, \widehat{\theta}_n, \alpha_n)$*, then* $\widehat{\psi}_n^{\text{sieve}}$ *is regular, asymptotically linear, and efficient for* $\Psi(P_0)$*, with influence function* χ_0*.*

For linear functionals and quadratic risks, the resulting autoDML estimators are doubly robust. They remain consistent if either $\|\alpha_n - \alpha_0\|_{\mathcal{H}} = o_p(1)$ or $\|\widehat{\theta}_n - \theta_0\|_{\mathcal{H}} = o_p(1)$, provided that $\|\eta_n - \eta_0\|_{\mathcal{N}} = o_p(1)$ and $(P_n - P_0)(\chi_n - \chi_0) = o_p(1)$. This extends the familiar double robustness of DML estimators for linear functionals of regression functions to a broad class of orthogonal losses, as in Example 1c. For such functionals, setting $\widehat{\theta}_n = 0$ gives the expected-gradient estimator $-P_n \dot{\ell}_{\eta_n}(0)(\alpha_n)$, which is consistent for $\psi_0(\theta_0)$ when $\|\alpha_n - \alpha_0\|_{\mathcal{H}} = o_p(1)$. This recovers inverse probability weighting and balancing estimators as special cases [Li et al., 2018, Ben-Michael et al., 2021].

The efficiency statement in Theorem 4.13 is for the loss-induced nonparametric projection parameter Ψ. As discussed in Section 4.3, this need not coincide with the sharp efficiency bound under a restricted model that treats $\mathcal{H}$ as a structural restriction. Under correct specification, the log-likelihood loss typically induces the projection whose nonparametric EIF coincides with the restricted-model EIF; more generally, χ_0 is an influence function for the restriction of Ψ to the restricted model, but is the restricted-model EIF only when it belongs to the corresponding tangent space.

For autoSieve, the required undersmoothing can be justified using cross-validation oracle inequalities [van der Laan and Dudoit, 2003, Laan et al., 2006]; see also Theorem 4 of Qiu et al. [2021]. Under mild assumptions, these results imply that the selected sieve $\mathcal{H}_{k(n)}$ approximates both θ_0 and α_0 through the projections $\theta_{0,k(n)}$ and $\alpha_{0,k(n)}$, so that Conditions B3 and B4 hold. Corollary 4.14 therefore yields an efficient debiased plug-in sieve estimator for orthogonal losses that may depend on estimated nuisances. This is useful when θ_0 is well suited to sieve estimation, while η_0 is more complex, higher-dimensional, or less smooth.

Ordinary finite-dimensional plug-in M-estimators are included as a special case. If $\mathcal{H}$ is finite-dimensional and $\theta_n := \operatorname{argmin}_{\theta \in \mathcal{H}} P_n \ell_{\eta_n}(Z, \theta)$, then $P_n m(\cdot, \theta_n)$ is the sieve estimator for the trivial sieve $\mathcal{H}_k = \mathcal{H}$. Corollary 4.14 implies that, under mild conditions, plug-in M-estimators based on Neyman-orthogonal losses with estimated nuisances are asymptotically linear and nonparametrically efficient for the associated projection parameter Ψ. In this finite-dimensional case, the limiting variance can be estimated by a sandwich estimator [White, 1982, Kauermann and Carroll, 2000] or by the bootstrap [Tang and Westling, 2024], using the loss ℓ_{η_n} and treating η_n as fixed.

4.5.2 Model misspecification error

In practice, the working model used by autoDML may be misspecified. Instead of estimating over the full space $\mathcal{H}$, we may restrict the M-estimand to a smaller closed linear space $\mathcal{H}' \subseteq \mathcal{H}$. The next theorem quantifies the resulting target error through a deterministic approximation formula. Appendix 4.E uses this comparison to study data-adaptive choices of $\mathcal{H}'$.

Given $\mathcal{H}'$, define

$$\theta_{0,\mathcal{H}'} := \operatorname*{argmin}_{\theta \in \mathcal{H}'} L_0(\theta, \eta_0), \qquad \Psi_{\mathcal{H}'}(P_0) := \psi_0(\theta_{0,\mathcal{H}'}).$$

We also write $\Psi_{\mathcal{H}}(P_0) := \psi_0(\theta_0)$. When it exists, we write $\alpha_{0,\mathcal{H}'}$ for the Hessian projection of α_0 onto $\mathcal{H}'$, that is,

$$\partial_\theta^2 L_0(\theta_0, \eta_0)(\alpha_0 - \alpha_{0,\mathcal{H}'}, h) = 0 \qquad \text{for all } h \in \mathcal{H}'.$$

Theorem 4.15 (Model approximation error)**.** *Suppose Conditions A1–A6 hold for $\mathcal{H}'$ as well as for $\mathcal{H}$, and that $\rho_{\mathcal{H}}(\alpha_{0,\mathcal{H}'}) < \infty$. Then*

$$\begin{aligned}\Psi_{\mathcal{H}'}(P_0) - \Psi_{\mathcal{H}}(P_0) = \partial_\theta^2 L_0(\theta_0, \eta_0)\big(\alpha_0 - \alpha_{0,\mathcal{H}'}, \theta_{0,\mathcal{H}'} - \theta_0\big) \\ + O\big(\|\theta_{0,\mathcal{H}'} - \theta_0\|_{\mathcal{H}}^2\big)\,,\end{aligned}$$

where the quadratic remainder vanishes when ψ_0 is linear and $L_0(\cdot, \eta_0)$ is quadratic at θ_0.

The same expansion yields an omitted-variable-bias formula for M-estimation. Suppose $\mathcal{H}$ is the class of functions of the full covariate set, while $\mathcal{H}'$ restricts attention to a subset of covariates. Then the bias is governed by the Hessian inner product between the component of θ_0 omitted by $\mathcal{H}'$ and the corresponding omitted component of the Riesz representer. When the risk is quadratic and the functional is linear, the remainder in Theorem 4.15 vanishes, so the bias is exactly this Hessian inner product. See Chernozhukov et al. [2021a] for the regression case and Section 3.3 of Ichimura and Newey [2022] for a related bound under exogeneity-based orthogonality conditions.

4.6 Numerical experiments

We use numerical experiments to illustrate how the proposed autoDML framework can be applied when the target is a nonlinear functional of a nonparametric M-estimand and the corresponding efficient influence function is not available in a simple closed form. Our main experiment studies inference on long-term survival probabilities under the beta-geometric survival model. This example is useful because the target depends on two covariate-dependent shape functions defined through a likelihood loss, and extrapolation beyond the observed follow-up horizon relies on the structure of the fitted survival model. Appendix 4.G.1 gives additional experiments for average treatment effect estimation based on the R-learner loss, a Neyman-orthogonal loss for the conditional average treatment effect [Nie and Wager, 2021].

4.6.1 The beta-geometric survival model

Let $T, C \in \mathbb{N}$ denote latent failure and censoring times, and let $Z = (X, \widetilde{T}, \Delta) \sim P_0$, where $\widetilde{T} = T \wedge C$ and $\Delta = \mathbb{1}(T \le C)$. The goal is to estimate the long-term survival probability $P_0(T > t_0)$ for a reference time t_0 that may lie beyond the observed follow-up horizon. Such extrapolation is important in applications such as subscription retention and manufacturing reliability, but is not identified by standard nonparametric survival methods without additional structure.

A simple structural model assumes $T \mid X \sim \text{Geometric}\{\pi(X)\}$, equivalently $P_0(T = t \mid T \ge t, X) = \pi(X)$. This imposes a time-homogeneous hazard, which is often too restrictive because failure risk typically changes with duration. The beta-geometric model [Hubbard et al., 2021] relaxes this restriction by introducing latent heterogeneity in the failure probability:

$$T \mid \Pi, X \sim \text{Geometric}(\Pi), \qquad \Pi \mid X \sim \text{Beta}\{\alpha_0(X), \beta_0(X)\},$$

where $\alpha_0(X) = \exp\{a_0(X)\}$ and $\beta_0(X) = \exp\{b_0(X)\}$. Marginalizing over Π yields the conditional hazard

$$\lambda_{a_0,b_0}(t \mid x) = \frac{\exp\{a_0(x)\}}{\exp\{a_0(x)\} + \exp\{b_0(x)\} + t - 1},$$

which decreases with t and can therefore accommodate higher failure risk at early durations. The parametric dependence on time enables extrapolation beyond the observed horizon, while the functions a_0 and b_0 allow flexible covariate heterogeneity.

Let $\mathcal{H} = \mathcal{H}_a \times \mathcal{H}_b$, with $\mathcal{H}_a, \mathcal{H}_b \subseteq L^2(\mu_X)$. The beta-geometric nuisance parameter is the M-estimand

$$\theta_0 = (a_0, b_0) = \operatorname*{argmin}_{(a,b)\in\mathcal{H}} E_0\{\ell(Z, (a, b))\},$$

where the negative log-likelihood loss is

$$\ell(z, (a, b)) = -\delta \log\{\lambda_{a,b}(t \mid x)\} - \sum_{s=1}^{t-1} \log\{1 - \lambda_{a,b}(s \mid x)\}.$$

The target can be written as the smooth functional

$$\psi_0(\theta_0) = E_0\{m(Z, \theta_0)\}, \qquad m(z, (a, b)) = \prod_{s=1}^{t_0} \{1 - \lambda_{a,b}(s \mid x)\}.$$

This example highlights the practical role of the proposed framework. Direct efficient estimation would require deriving the efficient influence function for a nonlinear functional of the likelihood-defined functions a_0 and b_0, which is cumbersome and may not yield a convenient closed-form representer. Theorem 4.11 instead characterizes the efficient influence function through the loss derivative and the Hessian Riesz representer. The autoDML estimator can therefore be implemented using only the gradient and Hessian of the beta-geometric loss and the derivative of the target functional, all of which follow by straightforward calculus; details are given in Appendix 4.F. Because the loss is the correctly specified negative log-likelihood, the resulting autoDML estimator is semiparametrically efficient under the beta-geometric model.

4.6.2 Experimental results

The data-generating distribution P_0, described in Appendix 4.F.3, includes three continuous covariates and one binary covariate. The failure time T follows a beta-geometric model with shape parameters a_0 and b_0, each specified as linear functions of the covariates. All observations are deterministically censored at $t = 6$. The target is the mean survival probability at $t_0 = 12$, which lies beyond the observable time horizon.

We estimate the M-estimand $\theta_0 = (a_0, b_0)$ using gradient-boosted trees with bivariate outputs, fitted by minimizing the negative log-likelihood loss ℓ. We estimate the Riesz representer $\alpha_0 \in \mathcal{H}$ using gradient-boosted trees trained on the corresponding Riesz loss. Both estimators are implemented in `xgboost` [Chen et al., 2015] with custom objectives, following Hubbard et al. [2021] for bivariate tree models. We impose an additive structure on $\mathcal{H} = \mathcal{H}_a \times \mathcal{H}_b$, with both $\mathcal{H}_a$ and $\mathcal{H}_b$ additive in the covariates; in `xgboost`, this is enforced through `interaction_constraints`.

We compare three debiased estimators of $\psi_0(\theta_0)$, all introduced in Section 4.4.1: two one-step autoDML estimators, with and without the TMLE-inspired stabilization in (4.7) and

Algorithm 8, and the autoTMLE estimator of Algorithm 9. As a benchmark, we also report the naive plug-in estimator. For simplicity, we use sample splitting: nuisance functions are estimated on an auxiliary sample, while point estimates and confidence intervals are computed on an independent main sample. For autoDML and autoTMLE, we construct 95% Wald confidence intervals using the estimated influence function.

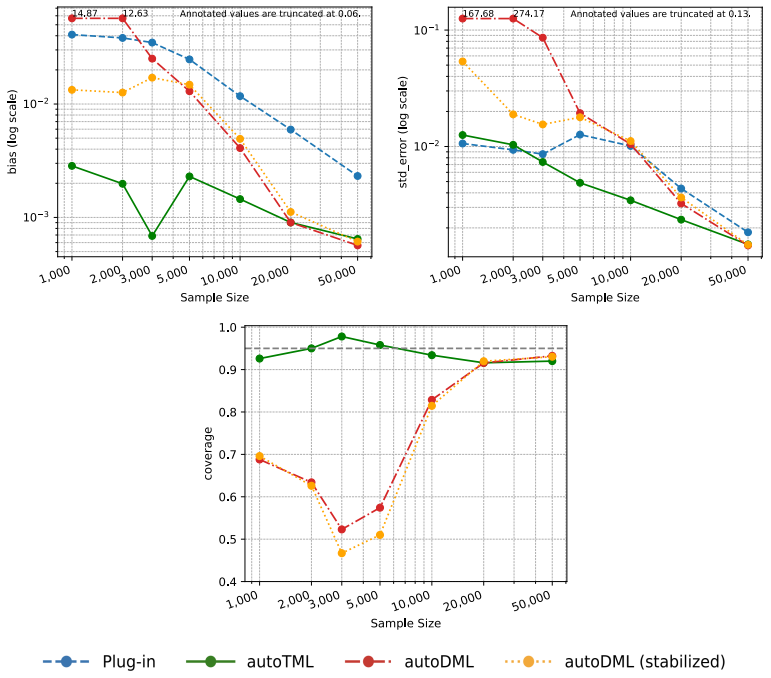


Figure 4.2: Monte Carlo performance as a function of sample size n: absolute bias (left), standard error (middle), and 95% confidence interval coverage (right).

Figure 4.2 reports Monte Carlo estimates of absolute bias, standard error, and 95% confidence interval coverage. For large sample sizes ($n \geq 20{,}000$), the three debiased estimators perform similarly, consistent with the asymptotic equivalence predicted by the theory. Their finite-sample behavior, however, differs substantially.

Among the debiased estimators, autoTMLE performs best throughout. It has the smallest bias and standard error across the full range of sample sizes, and its confidence interval coverage remains close to the nominal 95% level even in smaller samples. The stabilized autoDML estimator improves noticeably over the unstabilized version, with lower bias and greater stability, but it remains more biased and more variable than autoTMLE. The unstabilized autoDML estimator performs poorly in small samples: it is highly variable, exhibits substantial bias, and can perform worse than the naive plug-in estimator. Its coverage is also poor at smaller n, reflecting the magnitude of its bias.

These results suggest that, although autoDML and autoTMLE are asymptotically equivalent, their finite-sample performance can differ markedly. In this example, stabilization helps, but does not fully close the gap between autoDML and autoTMLE. A plausible ex-

planation is that the one-step correction used by autoDML depends on accurate local linear and quadratic approximations of the functional and loss around the initial M-estimator. When that initial estimator is not yet accurate, these approximations may be poor, limiting the effectiveness of the bias correction. By contrast, autoTMLE debiases through an explicit update of the M-estimand itself, after which inference is based on a plug-in estimator. This appears to yield greater robustness when the initial fit is imperfect.

4.7 Conclusion

We developed a general framework for debiased inference on smooth functionals of M-estimands defined by Neyman-orthogonal losses. The main message is that, in this setting, efficient influence-function derivations are modular. Given an orthogonal loss for the M-estimand, the efficient influence function is obtained from the derivative of the target, the derivative of the loss, and the Hessian Riesz representer linking the two. This representer determines both the leading plug-in bias and the correction needed to remove it, either explicitly through a one-step update or implicitly through a least favorable submodel for targeted minimum loss-based estimation. The same structure yields a common template for one-step, targeted, and sieve-based autoDML estimators.

Several components of the framework may be useful more broadly. First, Section 4.3 extends the least favorable submodel idea underlying targeted minimum loss-based estimation to general nuisance-dependent losses. Second, Section 4.4.1 introduces a stabilization rule for one-step autoDML estimators through (4.7). This normalization is particularly relevant for nonlinear targets and non-quadratic losses, and is consistent with the finite-sample behavior observed in our experiments. Third, Section 4.5.1 and Section 4.5.2 quantify sampling error and working-model approximation error for orthogonal M-estimation; Appendix 4.E extends this analysis to data-driven model selection.

The framework also admits several direct extensions. We focus on targets of the form $E_0\{m(Z,\theta)\}$ for a known map m, which covers many parameters in causal inference. Ordinary smooth functionals $\psi(\theta)$ are included by taking $m(Z,\theta) \equiv \psi(\theta)$. More general targets of the form $E_0\{m_{g_0}(Z,\theta)\}$, where g_0 is itself an M-estimand, can be handled by treating (θ_0, g_0) as a joint M-estimand; alternatively, one can start from an asymptotically linear estimator of the target functional [van der Laan et al., 2024c]. Although our development is stated for M-estimands, the same principle should extend to Z-estimands defined by infinite-dimensional Neyman-orthogonal estimating equations, given the close connection between M- and Z-estimation [van der Vaart and Wellner, 1996]. Appendix C.3 treats several such extensions, including an improved von Mises expansion under universal Neyman orthogonality, functionals of constrained M-estimands identified through a Lagrangian formulation, and the use of Hessian Riesz representers to orthogonalize moment equations.

Two practical questions deserve further study. First, for linear functionals and quadratic risks, autoDML estimators have a doubly robust expansion in the M-estimand and the Hessian representer. It would be useful to understand when this structure yields doubly robust inference, including valid confidence intervals [Benkeser et al., 2017, van der Laan et al., 2024c]. Second, Neyman-orthogonal losses are often nonconvex, which can complicate direct estimation of θ_0. Our framework allows θ_0 to be estimated by any suitable method,

including convex nonorthogonal losses, but automating the construction of computationally tractable orthogonal losses remains important. The EP-learning approach of van der Laan et al. [2024a], which constructs orthogonal risks from convex losses using debiased nuisance estimators, provides one promising route.

Chapter Appendix

4.A Cross-fitting and algorithms for autoDML

We now describe how to incorporate cross-fitting into the autoDML framework. Cross-fitting mitigates overfitting and relaxes empirical-process requirements for nuisance estimation [van der Laan et al., 2011, Chernozhukov et al., 2018a]. Here the main complication is that both θ_0 and α_0 depend on other nuisance functions, which themselves should be estimated out of fold. Algorithm 7 gives a computationally efficient procedure that cross-fits each nuisance only once. To estimate the Riesz representer, we assume that $\partial_\theta^2 L_0(\theta_0, \eta_0)(\alpha, \alpha) = P_0 \ddot{\ell}_{\eta_0}(\theta_0)(\alpha, \alpha)$ and $\dot{\psi}_0(\theta_0)(\alpha) = P_0 \dot{m}_{\theta_0}(\cdot, \alpha)$, so that α_0 can be learned from the pointwise Riesz loss

$$(z, \alpha) \mapsto \ddot{\ell}_{\eta_0}(\theta_0)(\alpha, \alpha)(z) - 2\, \dot{m}_{\theta_0}(z, \alpha).$$

Algorithm 7 Cross-fit nuisance estimation for autoDML

Input: dataset $\mathcal{D}_n = \{Z_i : i = 1, \ldots, n\}$, nuisance estimation algorithm $\mathcal{A}_\eta$, number J of folds

Partition $\mathcal{D}_n$ into folds $\mathcal{T}^{(1)}, \ldots, \mathcal{T}^{(J)}$. For each i, let $j(i)$ denote its fold and let $\mathcal{I}^{(s)} := \{i : Z_i \in \mathcal{T}^{(s)}\}$.
Cross-fit nuisance parameter.
for $s = 1, \ldots, J$ **do**
    set $\eta_{n,s} := \mathcal{A}_\eta(\mathcal{D}_n \setminus \mathcal{T}^{(s)})$.
end for
Cross-fit M-estimand.
for $s = 1, \ldots, J$ **do**
    set $\theta_{n,s}$ to the minimizer of $\theta \mapsto \sum_{i \notin \mathcal{I}^{(s)}} \ell_{\eta_{n,j(i)}}(Z_i, \theta)$.
end for
Cross-fit Hessian Riesz representer.
for $s = 1, \ldots, J$ **do**
    set $\alpha_{n,s}$ to the minimizer of $\alpha \mapsto \sum_{i \notin \mathcal{I}^{(s)}} \{\ddot{\ell}_{\eta_{n,j(i)}}(\theta_{n,j(i)})(\alpha, \alpha)(Z_i) - 2\, \dot{m}_{\theta_{n,j(i)}}(Z_i, \alpha)\}$.
end for
return cross-fitted estimates $\{\eta_{n,s}, \theta_{n,s}, \alpha_{n,s} : s = 1, \ldots, J\}$.

Algorithm 7 constructs cross-fitted estimates of η_0, θ_0, and α_0 in a single forward pass, with each stage using the previously cross-fitted nuisance estimates. This avoids the computational cost of re-cross-fitting all lower-level nuisances separately inside each training fold. The trade-off is a mild form of data reuse: the cross-fitted nuisance estimates used in the losses for θ_0 and α_0 are themselves constructed using observations from all folds, and the Riesz loss also uses cross-fitted estimates of θ_0. This leakage could be removed by nested cross-fitting, but at substantially higher computational cost. Similar reuse arises in Algorithm 3 of van der Laan et al. [2024a] for cross-validation with estimated nuisances and did not appear to affect empirical performance.

With the nuisance estimates from Algorithm 7, we define the cross-fitted autoDML and stabilized autoTMLE estimators in Algorithms 8 and 9. Both follow the debiasing template in Section 4.4.1, replacing each nuisance by its out-of-fold estimate in the final empirical average.

Algorithm 8 Cross-fitted one-step autoDML

Input: dataset $\mathcal{D}_n = \{Z_i : i = 1, \ldots, n\}$, nuisance estimation algorithm $\mathcal{A}_\eta$, number J of cross-fitting splits

`# Cross-fit nuisances.`

1: obtain cross-fitted estimators $\{\eta_{n,j}, \theta_{n,j}, \alpha_{n,j} : 1 \le j \le J\}$ from Alg. 7 with $\mathcal{D}_n$, $\mathcal{A}_\eta$, and J;

`# Optional TMLE-inspired stabilization via (4.7)`

2: for $s = 1, \ldots, J$, set $\alpha_{n,s} := \varepsilon_n \alpha_{n,s}$ with $\varepsilon_n := \frac{\sum_{i=1}^n \dot{m}_{\theta_{n,j(i)}}(Z_i, \alpha_{n,j(i)})}{\sum_{i=1}^n \ddot{\ell}_{\eta_{n,j(i)}}(\theta_{n,j(i)})(\alpha_{n,j(i)}, \alpha_{n,j(i)})(Z_i)}$.

`# Compute one-step estimator`

3: set $\widehat{\psi}_n^{\mathrm{dml}} := \frac{1}{n}\sum_{i=1}^n m(Z_i, \theta_{n,j(i)}) - \frac{1}{n}\sum_{i=1}^n \dot{\ell}_{\eta_{n,j(i)}}(\theta_{n,j(i)})(\alpha_{n,j(i)})(Z_i)$

4: **return** autoDML estimate $\widehat{\psi}_n^{\mathrm{dml}}$

Algorithm 9 Cross-fitted autoTMLE

Input: dataset $\mathcal{D}_n = \{Z_i : i = 1, \ldots, n\}$, nuisance estimation algorithm $\mathcal{A}_\eta$, number J of cross-fitting splits

`# Cross-fit nuisances.`

1: obtain cross-fitted estimators $\{\eta_{n,j}, \theta_{n,j}, \alpha_{n,j} : 1 \le j \le J\}$ from Alg. 7 with $\mathcal{D}_n$, $\mathcal{A}_\eta$, and J;

`# Targeting step.`

2: compute fluctuation $\varepsilon_n := \mathrm{argmin}_\varepsilon \sum_{i=1}^n \ell_{\eta_{n,j(i)}}(Z_i, \theta_{n,j(i)} + \varepsilon \alpha_{n,j(i)})$

3: for $s = 1, \ldots, J$, set $\theta^*_{n,s} := \theta_{n,s} + \varepsilon_n \alpha_{n,s}$;

`# Compute plug-in estimator.`

4: set $\widehat{\psi}_n^{\mathrm{tmle}} := \frac{1}{n}\sum_{i=1}^n m(Z_i, \theta^*_{n,j(i)})$

5: **return** autoTMLE estimate $\widehat{\psi}_n^{\mathrm{tmle}}$

4.B Connection to the canonical gradient in nonparametric models

We illustrate how the Hessian Riesz representer recovers the classical canonical gradient for a pathwise differentiable functional of a density. Let $\theta_0 := dP_0/d\mu$ denote the density of P_0 with respect to a dominating measure μ on $\mathcal{Z}$, and let $\mathcal{H} = L^2(\mu)$. Consider the least-squares risk and corresponding loss

$$L_0(\theta) = \frac{1}{2}\int \theta(z)^2 \, d\mu(z) - E_0[\theta(Z)], \quad \ell(z, \theta) = \frac{1}{2}\int \theta(u)^2 \, d\mu(u) - \theta(z).$$

We use the least-squares risk because it permits optimization over the unconstrained linear space $L^2(\mu)$, which aligns with the linear-space setup of our framework.

In this example, the Hessian induces the usual $L^2(\mu)$ inner product, so the Hessian Riesz representer α_0 of $\dot{\psi}(\theta_0)$ is simply its $L^2(\mu)$ Riesz representer:

$$\dot{\psi}(\theta_0)(h) = \int h(z)\alpha_0(z) \, d\mu(z) \qquad \text{for all } h \in \mathcal{H}.$$

For any direction $h \in L^\infty(\mu)$ satisfying $\int h(z) \, d\mu(z) = 0$, the submodel $(\theta_0 + th : t \in \mathbb{R})$ is locally a valid density submodel through θ_0 with score function $s_h = h/\theta_0$ at $t = 0$. Along

such submodels,

$$\dot{\psi}(\theta_0)(h) = E_0\big[s_h(Z)\{\alpha_0 - P_0\alpha_0\}(Z)\big] \qquad \text{for all } h \in \mathcal{H} \text{ such that } \int h(z)\, d\mu(z) = 0.$$

Thus, $\alpha_0 - P_0\alpha_0$ is the $L^2(P_0)$ Riesz representer, or canonical gradient, of the pathwise derivative map $s_h \mapsto \dot{\psi}(\theta_0)(h)$ on the tangent space $L_0^2(P_0)$ of mean-zero, finite-variance score functions [Bickel et al., 1993]. Under mild conditions, the classical tangent-space formulation [Bickel et al., 1993] implies that Ψ is pathwise differentiable, with efficient influence function

$$\chi_0(z) = \alpha_0(z) - E_0[\alpha_0(Z)] + m(z, \theta_0) - \Psi(P_0).$$

Our M-estimand framework recovers the same efficient influence function, since

$$-\dot{\ell}(\theta_0)(\alpha_0)(z) = \alpha_0(z) - \int \theta_0(u)\alpha_0(u)\, d\mu(u) = \alpha_0(z) - E_0[\alpha_0(Z)].$$

In the main theory, we allow the nuisance parameter to range over a convex class while retaining a linear ambient direction space for perturbations in η. By contrast, extending the primary parameter space $\mathcal{H}$ itself beyond the linear setting would require a separate tangent-cone or variational-inequality formulation. We therefore leave genuinely convex extensions of $\mathcal{H}$, such as boundary-constrained density classes under likelihood loss, for future work.

4.C Discussion of conditions

Condition B2 and Condition B3 impose that the estimators η_n and θ_n are, respectively, consistent for η_0 and θ_0 at a rate faster than $n^{-\frac{1}{4}}$. Such rate conditions are standard in semiparametric statistics and orthogonal machine learning and can be achieved under appropriate conditions by several algorithms, including generalized additive models [Hastie and Tibshirani, 1987], reproducing kernel Hilbert space estimators [Nie and Wager, 2021], the highly adaptive lasso [van der Laan, 2017], gradient-boosted trees [Schuler et al., 2023], and neural networks [Farrell et al., 2018]. Fast oracle rates for the estimator θ_n, which would be achieved if η_0 were known, can be obtained using orthogonal learning [Foster and Syrgkanis, 2023].

Condition B5 is a standard empirical process requirement that holds when $\|\chi_n - \chi_0\|_{L^2(P_0)} = o_p(1)$ and the realizations of $\chi_n - \chi_0$ belong to a fixed Donsker class. In debiased machine learning, sample-splitting and cross-fitting are commonly used to mitigate overfitting and relax Donsker conditions [van der Laan et al., 2011, Chernozhukov et al., 2018a]. In particular, B5 holds for arbitrary machine learning estimators if $\|\chi_n - \chi_0\|_{L^2(P_0)} = o_p(1)$ and the nuisance estimators η_n, θ_n, and α_n are trained on an external sample independent of $\{Z_i\}_{i=1}^n$. To improve data efficiency, one may instead use cross-fitting, as in Algorithms 7–9, where nuisance and parameter estimates are computed across multiple data splits. For our modified procedure in Algorithm 7, however, B5 does not follow from standard cross-fitting arguments due to data leakage across folds from reusing cross-fitted nuisance estimates. Nonetheless, we conjecture that it still holds under reasonable assumptions, since the leakage is limited. For example, the condition is satisfied if, for each $j \in [J]$, the realizations of the cross-fitted estimators $\{\theta_{n,j}, \alpha_{n,j}\}$ lie in a (possibly random) function class $\mathcal{F}_{n,j}$ that is conditionally fixed and uniformly Donsker [Sheehy and

Wellner, 1992], given the j-th training set.

4.C.1 Hellinger continuity of M-estimands

The Hellinger continuity of the M-estimand map $P \mapsto \theta_P$ in Condition A7 can often be verified directly from the population risk. The following lemma records a useful sufficient condition.

Lemma 4.16 (Hellinger continuity under local strong convexity)**.** *Let d_{H} denote Hellinger distance, and let $\|\cdot\|_{\mathcal{H}^*}$ denote the operator norm dual to $\|\cdot\|_{\mathcal{H}}$. For P in a Hellinger neighborhood*

$$\mathcal{P}_\delta := \{P \in \mathcal{P} : d_{\mathrm{H}}(P, P_0) < \delta\},$$

write $R_P(\theta) := L_P(\theta, \eta_P)$ and $\theta_P := \operatorname{argmin}_{\theta\in\mathcal{H}} R_P(\theta)$. Suppose that, for every $P \in \mathcal{P}_\delta$, the minimizer θ_P exists and R_P is Fréchet differentiable at θ_0. Assume further that there exist constants $\kappa > 0$ and $C_\nabla < \infty$ such that:

(i) R_P has a uniform quadratic lower bound at θ_0, in the sense that, for all $\theta \in \mathcal{H}$,

$$R_P(\theta) \geq R_P(\theta_0) + \partial_\theta R_P(\theta_0)(\theta - \theta_0) + \frac{\kappa}{2}\|\theta - \theta_0\|_{\mathcal{H}}^2.$$

(ii) The risk gradient at θ_0 is Hellinger Lipschitz:

$$\|\partial_\theta R_P(\theta_0) - \partial_\theta R_0(\theta_0)\|_{\mathcal{H}^*} \leq C_\nabla d_{\mathrm{H}}(P, P_0).$$

Then $P \mapsto \theta_P$ is Hellinger Lipschitz at P_0:

$$\|\theta_P - \theta_0\|_{\mathcal{H}} \leq \frac{2C_\nabla}{\kappa} d_{\mathrm{H}}(P, P_0), \qquad P \in \mathcal{P}_\delta.$$

Proof deferred to Appendix C.

The gradient condition in Lemma 4.16 can often be verified from primitive bounds. Let $f_{\eta,h} := \dot{\ell}_\eta(\theta_0)(h)$. By Condition A3i,

$$\{\partial_\theta R_P(\theta_0) - \partial_\theta R_0(\theta_0)\}(h) = (P - P_0) f_{\eta_0,h} + P\{f_{\eta_P,h} - f_{\eta_0,h}\}.$$

The first term is $O\{d_{\mathrm{H}}(P, P_0)\}$, uniformly over $\|h\|_{\mathcal{H}} \leq 1$, whenever

$$\sup_{P\in\mathcal{P}_\delta} \sup_{\|h\|_{\mathcal{H}}\leq 1} \{P f_{\eta_0,h}^2 + P_0 f_{\eta_0,h}^2\} < \infty,$$

by the Cauchy–Schwarz inequality and the standard Hellinger bound for differences of expectations. The second term is $O\{d_{\mathrm{H}}(P, P_0)\}$ if, for all $P \in \mathcal{P}_\delta$,

$$\sup_{\|h\|_{\mathcal{H}}\leq 1} \|f_{\eta_P,h} - f_{\eta_0,h}\|_{L^2(P)} \leq C\|\eta_P - \eta_0\|_{\mathcal{N}}$$

and the nuisance map $P \mapsto \eta_P$ is Hellinger Lipschitz at P_0. Thus, under these primitive bounds, the M-estimand continuity required in Condition A7 follows from the local curvature of the risk and Hellinger continuity of its gradient at θ_0.

4.D Examples

4.D.1 Remarks

Remark 4.17 (Separable product spaces and vector-valued M-estimands)**.** Suppose that $\mathcal{H} = \mathcal{H}_1 \times \cdots \times \mathcal{H}_K$, where each $\mathcal{H}_k$ is a Hilbert space, and write $\theta = (\theta_1, \ldots, \theta_K)$. Consider a loss of the form $\ell_{\eta_0}(z, \theta) = \sum_{k=1}^K \ell_{\eta_0}^{(k)}(z, \theta_k)$. Then the population risk is separable,

$$L_0(\theta, \eta_0) = \sum_{k=1}^K L_0^{(k)}(\theta_k, \eta_0),$$

where $L_0^{(k)}(\theta_k, \eta_0) := P_0 \ell_{\eta_0}^{(k)}(\cdot, \theta_k)$, so each component $\theta_{0,k}$ minimizes its own risk $L_0^{(k)}(\cdot, \eta_0)$. Moreover, for $h = (h_1, \ldots, h_K)$ and $h' = (h'_1, \ldots, h'_K)$,

$$\partial_\theta^2 L_0(\theta_0, \eta_0)(h, h') = \sum_{k=1}^K \partial_{\theta_k}^2 L_0^{(k)}(\theta_{0,k}, \eta_0)(h_k, h'_k).$$

It follows that the Hessian Riesz representer in the product space $\mathcal{H}$ decomposes componentwise as

$$\alpha_{0,\mathcal{H}} = (\alpha_{0,\mathcal{H}_1}^{(1)}, \ldots, \alpha_{0,\mathcal{H}_K}^{(K)}),$$

where $\alpha_{0,\mathcal{H}_k}^{(k)}$ is the Riesz representer of the map $h_k \mapsto \partial_\theta \psi_0(\theta_0)(0, \ldots, 0, h_k, 0, \ldots, 0)$ with respect to the k-th component Hessian inner product $(h_k, h'_k) \mapsto \partial_{\theta_k}^2 L_0^{(k)}(\theta_{0,k}, \eta_0)(h_k, h'_k)$. Consequently,

$$\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0) = \sum_{k=1}^K \dot{\ell}_{\eta_0}^{(k)}(\theta_{0,k})(\alpha_{0,\mathcal{H}_k}^{(k)}).$$

4.D.2 Loss functions satisfying conditions

Lemma 4.7 verifies that Conditions A3 and A4 hold for a broad class of twice-differentiable loss functions, but such smoothness is not strictly necessary. In particular, many common losses—such as the quantile loss—do not satisfy classical second-order smoothness, yet still induce a twice-differentiable risk functional.

Example 3 (Smoothness of median loss)**.** Suppose $Z = (X, Y)$ with $X \in \mathbb{R}^d$ and $Y \in \mathbb{R}$, and consider the median loss $\ell(\theta, z) = |y - \theta(x)|$. This loss is almost everywhere differentiable, with a generalized second derivative that exists in the distributional sense. Under standard regularity conditions on the conditional density $p_0(Y \mid X)$ (see Appendix 4.D.4), the population risk admits a bounded and Lipschitz continuous second derivative, $\partial_\theta^2 L_0(\theta, \eta)(h_1, h_2) = \int p_0(Y = \theta(x) \mid X = x) \cdot h_1(x) h_2(x) \, P_0(dx)$. Hence, A3 is satisfied for the median loss, and more generally for orthogonal quantile losses, such as those arising in quantile treatment effect settings [Leqi and Kennedy, 2021, Whitehouse et al., 2024]. □

Continuing with our examples in Section 4.2.3, we demonstrate how various functionals and loss functions of interest in causal inference satisfy our conditions. In each of the following examples, suppose each $P \in \mathcal{P}$ is dominated by P_0 and satisfies $c\|\cdot\|_{L^2(P_0)} \leq \|\cdot\|_{L^2(P)} \leq C\|\cdot\|_{L^2(P_0)}$ for constants $0 < c < C < \infty$. Let $\mathcal{H} = L^2(P_0)$, with $\|\cdot\|_{\mathcal{H}}$ the $L^2(P_0)$ norm and $\rho_{\mathcal{H}}(\cdot)$ the P_0-essential supremum norm.

Example 1b (continued)**.** Let $\eta_0 := (\pi_0, \mu_0(1,\cdot), \mu_0(0,\cdot))$ denote the nuisance parameter. Let $\mathcal{N} = L^\infty(\lambda) \times L^\infty(\lambda) \times L^\infty(\lambda)$. The orthogonal loss for semiparametric logistic regression can be written as $\ell_{\eta_0}(\theta, z) = l(\pi_0(x), \mu_0(0,x), \mu_0(1,x), \theta(x), z)$, where the function $l : \mathbb{R}^4 \times \mathcal{Z} \to \mathbb{R}$ is defined by

$$l(\overline{\pi}, \overline{\mu}^0, \overline{\mu}^1, \overline{\theta}, z) := g_1(a, \overline{\mu}^0, \overline{\mu}^1) \left\{ \log\left(1 + \exp\left(-(a - \overline{\pi})\,\overline{\theta} - g_2(\overline{\pi}, \overline{\mu}^0, \overline{\mu}^1)\right)\right) - y(a - \overline{\pi})\,\overline{\theta} \right\},$$

where $g_1(a, \overline{\mu}^0, \overline{\mu}^1) = \{(1-a)\overline{\mu}^0 + a\overline{\mu}^1\}^{-1}$ and $g_2(\overline{\pi}, \overline{\mu}^0, \overline{\mu}^1) = (1 - \overline{\pi}) \operatorname{logit}(\overline{\mu}^0) + \overline{\pi} \operatorname{logit}(\overline{\mu}^1)$. Under reasonable conditions, the function l satisfies the assumptions of Lemma 4.7, and Conditions A3 and A4 hold (see Appendix 4.D.4 for details). In particular, A3i is satisfied at P_0, with the map $h \mapsto \dot{\ell}_{\eta_0}(\theta)(h)$ given elementwise by $z \mapsto -h(x) \cdot \frac{a - \pi_0(x)}{\nu_0(a,x)} [y - \operatorname{expit}\{(a - \pi_0(x))\theta(x) + h_0(x)\}]$. Furthermore, A3ii is satisfied with $\partial_\theta^2 L_0(\theta, \eta_0)(h_1, h_2) = E_0\left[\{A - \pi_0(X)\}^2 h_1(X) h_2(X)\right]$. A6 holds if $1 - \delta > \pi_0(X) > \delta$ P_0-almost surely for some $\delta > 0$. Finally, when the semiparametric logistic regression model is correctly specified at P_0, A5 holds by Nekipelov et al. [2022] (see Equation (2.24) and Remark 2.2). The EIF term $-\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0)$ in Theorem 4.11 equals $z \mapsto \alpha_0(x) \frac{(a - \pi_0(x))}{\nu_0(a,x)} [y - \operatorname{expit}\{(a - \pi_0(x))\theta_0(x) + h_0(x)\}]$, where $\alpha_0 = \operatorname{argmin}_{\alpha \in \mathcal{H}} E_0[\pi_0(X)\{1 - \pi_0(X)\}\alpha^2(X) - 2\,\dot{m}_{\theta_0}(Z, \alpha)]$ is a weighted Riesz representer with overlap weights $\pi_0\{1 - \pi_0\}$. □

Example 1c (continued)**.** Consider the pseudo squared-error loss $\ell_\eta : (z, \theta) \mapsto \frac{1}{2} w_\eta(z) \{\zeta_\eta(z) - \theta(x)\}^2$. Suppose $\operatorname{ess\,sup}_{z \in \mathcal{Z}, P \in \mathcal{P}} \max\{w_{\eta_P}(z), \zeta_{\eta_P}(z), \theta_P(x)\} < \infty$ with respect to the dominating measure λ. This loss satisfies A3i with $\dot{\ell}_{\eta_0}(\theta)(h) : z \mapsto -w_{\eta_0}(z) h(x) \{\zeta_{\eta_0}(z) - \theta(x)\}$, and A3ii with $\partial_\theta^2 L_0(\theta, \eta_0)(h_1, h_2) = E_0[w_{\eta_0}(Z) h_1(X) h_2(X)]$. Conditions A3iii and A4 can be verified for many loss functions by applying Lemma 4.7. Condition A6 holds if $w_{\eta_0}(Z) > \delta$ P_0-almost surely for some $\delta > 0$. For the R-learner loss, recall that $w_{\eta_0} : z \mapsto (a - \pi_0(x))^2$ and $\zeta_{\eta_0} : z \mapsto \frac{y - m_0(x)}{a - \pi_0(x)}$, where $\eta_0 = (\pi_0, m_0)$ with $m_0 : x \mapsto E_0[Y \mid X = x]$. We can compute the cross derivative to be

$$\partial_\eta \partial_\theta L_0(\theta, \eta_0)\big((\delta\pi, \delta m), h\big) = E_0\Big[\Big\{ -(A - \pi_0(X))\,\delta\pi(X)\,\big(\zeta_{\eta_0}(Z) - \theta(X)\big) \\ - (A - \pi_0(X))\,\delta m(X) + (Y - m_0(X))\,\delta\pi(X) \Big\}\, h(X)\Big].$$

The first term on the right-hand side equals $E_0[\,\delta\pi(X)\,\big(Y - m_0(X) - (A - \pi_0(X))\theta(X)\big)]$, which is zero since $E_0[Y - m_0(X) \mid X] = 0$ and $E_0[A - \pi_0(X) \mid X] = 0$. The second term also vanishes for the same reason. Hence, A5 holds, and, in fact, the R-learner loss is universally orthogonal, satisfying $\partial_\eta \partial_\theta L_0(\theta, \eta_0) = 0$ for all $\theta \in \mathcal{H}$. □

4.D.3 Examples of estimators

Example 1b (continued)**.** Consider the orthogonal semiparametric logistic regression loss, and let θ_n and α_n be estimators of θ_0 and α_0, respectively, where $\alpha_0 = \operatorname{argmin}_{\alpha \in \overline{\mathcal{H}}} E_0[\pi_0(X)\{1 - \pi_0(X)\}\alpha^2(X) - 2\,\dot{m}_{\theta_0}(Z, \alpha)]$. Let π_n be an estimator of the propensity score π_0, and let μ_n be an estimator of the outcome regression μ_0. Define $\nu_n = \mu_n(1 - \mu_n)$ and $h_n(x) = \pi_n(x) \operatorname{logit} \mu_n(1, x) + (1 - \pi_n(x)) \operatorname{logit} \mu_n(0, x)$. Then, an

autoDML estimator of $\psi_0(\theta_0)$ is given by

$$\frac{1}{n}\sum_{i=1}^{n} m(Z_i,\theta_n)+\frac{1}{n}\sum_{i=1}^{n}\alpha_n(X_i)\frac{(A_i-\pi_0(X_i))}{\nu_n(A_i,X_i)}\left[y_i-\text{expit}\{(A_i-\pi_n(X_i))\theta_n(X_i)+h_n(X_i)\}\right]. \square$$

Example 1c (continued)**.** For general pseudo-outcome regressions, an autoDML estimator is given by $\widehat{\psi}_n^{\text{dml}} := \frac{1}{n}\sum_{i=1}^n m(Z_i,\theta_n)+\frac{1}{n}\sum_{i=1}^n w_{\eta_n}(Z_i)\alpha_n(X_i)\{\zeta_{\eta_n}(Z_i)-\theta_n(X_i)\}$. For the R-learner loss of the CATE $\theta_0 : x \mapsto \mu_0(1,x)-\mu_0(0,x)$, corresponding to $w_{\eta_0} : z \mapsto (a-\pi_0(x))^2$ and $\zeta_{\eta_0} : z \mapsto \frac{y-m_0(x)}{a-\pi_0(x)}$, we have $-\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0) : z \mapsto \alpha_0(x)\{a-\pi_0(x)\}\{y-m_0(x)-(a-\pi_0(x))\theta_0(x)\}$, where $\alpha_0 = \text{argmin}_{\alpha\in\mathcal{H}} E_0[\pi_0(X)\{1-\pi_0(X)\}\alpha^2(X)]-2\dot{\psi}_0(\theta_0)(\alpha)$ is an overlap-weighted Riesz representer. Taking ψ_0 as the ATE, with $\psi_0(\alpha) = \dot{\psi}_0(\theta_0)(\alpha) = E_0[\alpha(X)]$, we find that α_0 is the overlap-weighted projection of $\{\pi_0(x)(1-\pi_0(x))\}^{-1}$ onto the CATE model $\mathcal{H}$. An autoDML estimator of the ATE is given by

$$\widehat{\psi}_n^{\text{dml}} := \frac{1}{n}\sum_{i=1}^{n}\theta_n(Z_i)+\frac{1}{n}\sum_{i=1}^{n}(A_i-\pi_n(X_i))\alpha_n(X_i)\{Y_i-m_n(X_i)-(A_i-\pi_n(X_i))\theta_n(X_i)\},$$

where $\eta_n := (\pi_n, m_n)$ is an estimator of $\eta_0 := (\pi_0, m_0)$, θ_n is an estimator of the CATE, and α_n is an estimator of α_0. For the R-learner loss, the fluctuation parameter for autoTMLE is $\varepsilon_n = \text{argmin}_\varepsilon\{Y_i - m_n(X_i) - \{A_i-\pi_n(X_i)\}\{\theta_n(X_i)+\varepsilon\alpha_n(X_i)\}\}^2$, and the autoTMLE estimator for the ATE is

$$\widehat{\psi}_n^{\text{tmle}} = \frac{1}{n}\sum_{i=1}^{n}\theta_n(X_i)+\frac{\sum_{i=1}^n \alpha_n(X_i)}{\sum_{i=1}^n (A_i-\pi_n(X_i))^2\,\alpha_n(X_i)^2}\frac{1}{n}\sum_{i=1}^{n}(A_i-\pi_n(X_i))\alpha_n(X_i)\{Y_i-m_n(X_i)-(A_i-\pi_n(X_i))\theta_n(X_i)\}. \square$$

4.D.4 Additional technical details for examples

In each of the following examples, suppose that each $P \in \mathcal{P}$ is dominated by a measure λ, and that $c\|\cdot\|_{L^2(\lambda)} \le \|\cdot\|_{L^2(P)} \le C\|\cdot\|_{L^2(\lambda)}$ for some $0 < c < C < \infty$. Let $\mathcal{H} = L^2(\lambda)$, $\mathcal{H} = L^\infty(\lambda)$, $\|\cdot\|_{\mathcal{H}}$ be the $L^2(P)$ norm, and $\rho_{\mathcal{H}}(\cdot)$ be the λ-essential supremum norm. By norm equivalence, $\mathcal{H}$ is closed in $L^2(P)$ for each $P \in \mathcal{P}$.

Example 1c (Smoothness of median loss)**.** Suppose $Z = (X, Y)$ with $X \in \mathbb{R}^d$ and $Y \in \mathbb{R}$, and consider the median loss $\ell(\theta, z) = |y - \theta(x)|$. Then ℓ is almost everywhere differentiable with $\dot{\ell}(\theta)(h)(z) = \text{sign}(y-\theta(x))\cdot h(x)$. The second derivative exists only in a distributional sense and is given by $\ddot{\ell}(\theta)(h_1,h_2)(z) = \delta(y-\theta(x))\cdot h_1(x)h_2(x)$, where δ is the Dirac delta function. Consequently, the second derivative of the population risk satisfies $\partial_\theta^2 L_0(\theta,\eta)(h_1,h_2) = \int p_0(Y=\theta(x)\mid X=x)\cdot h_1(x)h_2(x)\,P_0(dx)$. Let $\mathcal{H} := L^\infty(\lambda)$ with $\rho_{\mathcal{H}}(\cdot) = \|\cdot\|_{L^\infty(\lambda)}$ and $\|\cdot\|_{\mathcal{H}} = \|\cdot\|_{L^2(P_0)}$. Suppose $\text{ess sup}_{x\in\mathcal{X},\,\theta\in\mathcal{H}_{\mathcal{P}},P\in\mathcal{P}}\, p(Y=\theta(x)\mid X=x) < \infty$, and that the map $y \mapsto p(Y=y\mid X=x)$ is Lipschitz continuous for each $P \in \mathcal{P}$ with constant $L_x < \infty$, where $L := \text{ess sup}_x L_x < \infty$. Then, uniformly over all $h_1, h_2 \in \mathcal{H}$, the second derivative is bounded, satisfying $\partial_\theta^2 L_0(\theta,\eta)(h_1,h_2) \le C\|h_1\|_{\mathcal{H}}\|h_2\|_{\mathcal{H}}$, and Lipschitz continuous with

$$|\partial_\theta^2 L_0(\theta+h,\eta)(h_1,h_2) - \partial_\theta^2 L_0(\theta,\eta)(h_1,h_2)| \le \int L_x|h(x)||h_1(x)h_2(x)|\,P_0(dx) \le L\rho_{\mathcal{H}}(h_1)\|h_2\|_{\mathcal{H}}\|h\|_{\mathcal{H}}.$$

Hence, Condition A3 is satisfied for the median loss, and more generally for orthogonal quantile losses, such as those arising in quantile treatment effect estimation [Leqi and Kennedy, 2021]. $\square$

Example 1c (Verifying conditions for semiparametric orthogonal logistic loss)**.** Let $\eta_0 := (\pi_0, \mu_0(1,\cdot), \mu_0(0,\cdot))$ denote the nuisance parameter. Let $\mathcal{N} = L^\infty(\lambda) \times L^\infty(\lambda) \times L^\infty(\lambda)$. We identify functions of subvectors of $z \in \mathcal{Z}$ with elements of $L^\infty(\lambda)$ via composition with coordinate projections; for example, $f(z) := f(x)$ when $z = (x, y)$. The orthogonal loss for semiparametric logistic regression can be written as $\ell_{\eta_0}(\theta, z) = l(\pi_0(x), \mu_0(0,x), \mu_0(1,x), \theta(x), z)$, where the function $l : \mathbb{R}^4 \times \mathcal{Z} \to \mathbb{R}$ is defined by

$$l(\overline{\pi}, \overline{\mu}^0, \overline{\mu}^1, \overline{\theta}, z) := g_1(a, \overline{\mu}^0, \overline{\mu}^1) \left\{ \log\left(1 + \exp\left(-(a - \overline{\pi})\,\overline{\theta} - g_2(\overline{\pi}, \overline{\mu}^0, \overline{\mu}^1)\right)\right) - y(a - \overline{\pi})\,\overline{\theta} \right\},$$

where $g_1(a, \overline{\mu}^0, \overline{\mu}^1) = \{(1-a)\overline{\mu}^0 + a\overline{\mu}^1\}^{-1}$ and $g_2(\overline{\pi}, \overline{\mu}^0, \overline{\mu}^1) = (1 - \overline{\pi})\,\mathrm{logit}(\overline{\mu}^0) + \overline{\pi}\,\mathrm{logit}(\overline{\mu}^1)$. Suppose that for all $P \in \mathcal{P}$ and $z \in \mathcal{Z}$, we have $\delta < \mu_P(a,x) < 1 - \delta$ and $\max\{|h_P(x)|, |\theta_P(x)|\} < M'$ almost surely, for some $\delta \in (0,1)$ and $M' < \infty$, with $\mathcal{Z}$ compact. Then A8 holds and the closure of the set $\{(\pi(x), \mu(0,x), \mu(1,x), \theta(x), z) : \theta \in \mathcal{H}_K, (\pi, \mu(0,\cdot), \mu(1,\cdot)) \in \mathcal{N}_K, z \in \mathcal{Z}\}$ is compact. Moreover, the function l is three times Lipschitz continuously differentiable in all arguments on this compact set, where we use that $t \mapsto t^{-1}$ on a closed subset of $(0,1)$ and logit on $[-M', M']$ are infinitely often Lipschitz continuously differentiable. Hence, l satisfies the conditions of Lemma 4.7, and Conditions A3 and A4 hold. In particular, A3i is satisfied at P_0, with the map $h \mapsto \dot{\ell}_{\eta_0}(\theta)(h)$ given elementwise by $z \mapsto -h(x) \cdot \frac{a - \pi_0(x)}{\nu_0(a,x)} \left[y - \mathrm{expit}\left\{(a - \pi_0(x))\theta(x) + h_0(x)\right\}\right]$, where $\|\cdot\|_{\mathcal{H}} := \|\cdot\|_{L^2(P_0)}$. Furthermore, A3ii is satisfied with $\partial_\theta^2 L_0(\theta, \eta_0)(h_1, h_2) = E_0\left[\{A - \pi_0(X)\}^2 h_1(X) h_2(X)\right]$. Condition A6 holds if $1 - \delta > \pi_0(X) > \delta$ P_0-almost surely for some $\delta > 0$. Condition A1 holds if $1 - \delta > \pi_P(X) > \delta$ uniformly over $P \in \mathcal{P}$, in which case the risk $\theta \mapsto L_P(\theta, \eta_P)$ is strongly convex and continuous on $\mathcal{H}$. Finally, when the semiparametric logistic regression model is correctly specified at P_0, A5 holds by Nekipelov et al. [2022] (see Equation (2.24) and Remark 2.2). □

4.E Model selection with Adaptive Debiased Machine Learning

4.E.1 Data-driven working models

Adaptive Debiased Machine Learning (ADML) [van der Laan et al., 2023] combines debiased machine learning with data-driven model selection to construct superefficient, adaptive estimators of smooth functionals, allowing model assumptions to be learned from the data. The key idea is to estimate $\Psi_n(P_0)$ for a data-adaptive, projection-based parameter $\Psi_n : \mathcal{P} \to \mathbb{R}$, defined through a submodel $\mathcal{P}_n \subseteq \mathcal{P}$ learned from the data. If $\mathcal{P}_n$ stabilizes asymptotically to a fixed oracle submodel $\mathcal{P}_0 := \mathcal{P}_{P_0}$ in a suitable sense, the resulting estimator remains efficient for an oracle target $\Psi_0 : \mathcal{P} \to \mathbb{R}$ defined through $\mathcal{P}_0$. This oracle parameter coincides with the true target Ψ on $\mathcal{P}_0$, so that $\Psi_0(P_0) = \Psi(P_0)$ when $P_0 \in \mathcal{P}_0$. Crucially, Ψ_0 may have a substantially smaller efficiency bound at P_0 than Ψ, leading to more efficient estimators and tighter confidence intervals, while preserving unbiasedness. The simplest example is when $\mathcal{P}_n := \mathcal{P}_{k(n)}$ with $k(n) \to \infty$, obtained via model selection over a sieve $\mathcal{P}_1 \subseteq \mathcal{P}_2 \subseteq \cdots \subseteq \mathcal{P}$, and the limiting model $\mathcal{P}_0$ is either the full model $\mathcal{P}$ or some element of the sequence containing P_0.

We apply the ADML framework to construct adaptive estimators of $\Psi(P_0)$ by selecting a data-driven working model for the M-estimand θ_0. Let $\mathcal{H}_n \subseteq \mathcal{H}$ denote this model, and suppose it approximates an unknown oracle submodel $\mathcal{H}_0 \subseteq \mathcal{H}$. The corresponding

statistical models are $\mathcal{P}_n := \{P \in \mathcal{P} : \theta_P \in \mathcal{H}_n\}$ and $\mathcal{P}_0 := \{P \in \mathcal{P} : \theta_P \in \mathcal{H}_0\}$. We define the working and oracle parameters by $\Psi_{\mathcal{H}_n}(P) := \psi_P(\theta_{P,\mathcal{H}_n})$ and $\Psi_{\mathcal{H}_0}(P) := \psi_P(\theta_{P,\mathcal{H}_0})$, where $\theta_{P,\mathcal{H}} := \text{argmin}_{\theta\in\mathcal{H}} L_P(\theta, \eta_P)$. Let $\chi_{P,\mathcal{H}_n}$ and $\chi_{P,\mathcal{H}_0}$ denote the efficient influence functions of $\Psi_{\mathcal{H}_n}$ and $\Psi_{\mathcal{H}_0}$, respectively, as given in Theorem 4.11. For example, $\mathcal{H}_n$ may be selected via cross-validation over a sieve of models $\mathcal{H}_1 \subset \mathcal{H}_2 \subset \cdots \subset \mathcal{H}_\infty := \mathcal{H}$, where $\mathcal{H}$ is a correctly specified model containing θ_0, and $\mathcal{H}_0$ is the smallest correctly specified model in the sieve. Alternatively, $\mathcal{H}_n$ could be obtained via variable selection or a data-driven feature transformation, with $\mathcal{H}_0$ corresponding to the limiting set of selected variables or the limiting feature transformation.

Given the selected model $\mathcal{H}_n$, our proposed ADML estimator of $\Psi(P_0)$ is the debiased estimator $\widehat{\psi}_{n,\mathcal{H}_n}$ of the data-adaptive working parameter $\Psi_{\mathcal{H}_n} : \mathcal{P} \to \mathbb{R}$, constructed using any method from Section 4.4.1. We show that, under suitable conditions, this estimator remains valid even if $\mathcal{H}_n$ is misspecified, provided the model approximation error vanishes asymptotically. A key step is showing that the parameter approximation bias $\Psi_{\mathcal{H}_n}(P_0) - \Psi_{\mathcal{H}_0}(P_0)$ is second-order in the model error and thus asymptotically negligible. Conditionally on the selected spaces, Theorem 4.15 applied with working space $\mathcal{H}_n$ inside the comparison space $\mathcal{H}_{n,0} := \mathcal{H}_n + \mathcal{H}_0$, using closures if needed, gives this second-order behavior.

We now present our main result on the asymptotic linearity and superefficiency of the ADML estimator $\widehat{\psi}_{n,\mathcal{H}_n}$ for $\Psi(P_0)$. To establish this result, we assume the following conditions.

(C1) *Linear expansion:* $\widehat{\psi}_{n,\mathcal{H}_n} - \Psi_n(P_0) = (P_n - P_0)\chi_{0,\mathcal{H}_n} + o_p(n^{-1/2})$.

(C2) *Stabilization of selected model:* $n^{1/2}(P_n - P_0)\{\chi_{0,\mathcal{H}_n} - \chi_{0,\mathcal{H}_0}\} = o_p(1)$.

(C3) *Target approximate rate:* If the functional is not linear or the risk is not quadratic, then $\|\theta_{0,\mathcal{H}_n} - \theta_0\|_{\mathcal{H}} = o_p(n^{-\frac{1}{4}})$.

(C4) *Doubly robust approximate rate:* $\partial^2_\theta L_0(\theta_0, \eta_0)(\alpha_{0,\mathcal{H}_{n,0}} - \alpha_{0,\mathcal{H}_n}, \theta_{0,\mathcal{H}_n} - \theta_0) = o_p(n^{-1/2})$.

Theorem 4.4. *Assume that Conditions A1-A8 hold for the model $\mathcal{H}_{n,0}$. Suppose that $\mathcal{H}_n$ converges to an oracle submodel $\mathcal{H}_0$ with $\theta_0 \in \mathcal{H}_0$ in the sense that Conditions C1-C4 hold. Then, $\widehat{\psi}_{n,\mathcal{H}_n} - \Psi(P_0) = P_n\chi_{0,\mathcal{H}_0} + o_p(n^{-1/2})$, and $\widehat{\psi}_{n,\mathcal{H}_n}$ is a locally regular and efficient estimator for the oracle parameter Ψ_0 under the nonparametric statistical model.*

It is interesting to apply Theorem 4.4 with $\mathcal{H}_n = \mathcal{H}_{k(n)}$ and $\mathcal{H}_0 = \mathcal{H}$, where $\mathcal{H}_{k(n)}$ is an element of the sieve $\mathcal{H}_1 \subset \mathcal{H}_2 \subset \mathcal{H}_3 \subset \cdots \subset H_\infty := \mathcal{H}$ with $k(n) \to \infty$. In this case, the theorem establishes the asymptotic linearity and efficiency of sieve-based plug-in estimators based on orthogonal losses, generalizing and extending the results of Shen [1997], Chen and Liao [2014] and Hahn et al. [2018].

Condition C1 requires that the ADML estimator $\widehat{\psi}_{n,\mathcal{H}_n}$ is debiased for the data-adaptive working parameter Ψ_n, which can be established under conditions analogous to Theorem 4.13 applied with $\mathcal{H} := \mathcal{H}_n$. Conditions C2–C4, drawn from prior ADML work [van der Laan et al., 2023, 2026], ensure that data-driven model selection does not invalidate the debiased estimator. Condition C2 is an asymptotic stability condition requiring

that the learned model $\mathcal{H}_n$ converges to a fixed oracle submodel $\mathcal{H}_0$, typically formalized as $\|\chi_{0,\mathcal{H}_n} - \chi_{0,\mathcal{H}_0}\|_{L^2(P_0)} = o_p(1)$. This holds when the nuisance functions $\alpha_{0,\mathcal{H}_n}$ and $\theta_{0,\mathcal{H}_n}$ converge to their oracle counterparts $\alpha_{0,\mathcal{H}_0}$ and θ_0 in an appropriate sense. Conditions C3 and C4 further ensure that the approximation bias from using $\mathcal{H}_n$ instead of $\mathcal{H}_0$ is negligible, by requiring sufficiently fast convergence of $\alpha_{0,\mathcal{H}_n}$ and $\theta_{0,\mathcal{H}_n}$. Combined with Theorem 4.15, these conditions imply that $\Psi_n(P_0) - \Psi(P_0) = o_p(n^{-1/2})$.

4.F autoDML for beta-geometric survival model

4.F.1 Functional and loss

Following our autoDML framework, for $P \in \mathcal{P}$, we define our M-estimand θ_P as the tuple (a_P, b_P) and our target space as the product Hilbert space $\mathcal{H} := \mathcal{H}_a \times \mathcal{H}_b$, where $\mathcal{H}_a \subseteq L^2(\mu_X)$ and $\mathcal{H}_b \subseteq L^2(\mu_X)$ for some measure μ_X on $\text{supp}(X)$. Under the conditional ignorability of the censoring mechanism, the log-transformed shape parameters $\theta_P = (a_P, b_P)$ are identified via the following optimization problem:

$$\theta_P = \underset{(a,b)\in\mathcal{H}}{\text{argmin}}\, E_P[\ell(Z, a, b)],$$

where ℓ is the negative log-likelihood loss given by:

$$\ell : (z, a, b) \mapsto -\delta \log\big(\lambda_{a,b}(t \mid x)\big) - \sum_{s=1}^{t-1} \log\big(1 - \lambda_{a,b}(s \mid x)\big),$$

where we use $z := (x, t, \delta)$ to denote a realization of Z. Given a reference time $t_0 > 0$, we define our target parameter as $\psi_P(\theta_P) = E_P[m(Z, \theta_P)]$, where the functional $m : \mathcal{Z} \times \mathcal{H} \to \mathbb{R}$ is given by:

$$m(z, a, b) := \prod_{s=1}^{t_0} \big(1 - \lambda_{a,b}(s \mid x)\big),$$

which represents the mean survival probability $P(T > t_0)$ under the beta-geometric model. In Appendix 4.F.2, we derive the first and second derivatives $\dot{\ell}(\theta)$ and $\ddot{\ell}(\theta)$ of the loss ℓ at $\theta \in \mathcal{H}$, as well as the first derivative $\dot{m}_\theta$ of the functional m, which together determine the loss function for the Riesz representer $\alpha_0 \in \mathcal{H}$ in Theorem 4.11. Since we use the negative log-likelihood loss, as discussed below Theorem 4.13, the autoDML estimator is not only nonparametrically efficient for the projection parameter Ψ, but also semiparametrically efficient for $P_0(T > t_0)$ under the beta-geometric model.

4.F.2 Derivation of gradients and Hessians

It is shown in Hubbard et al. [2021] that the conditional distribution of T given X under a beta-geometric distribution P is determined by the following recursions:

$$\begin{aligned}
P(T = 1 \mid X) &= \frac{\alpha_P(X)}{\alpha_P(X) + \beta_P(X)}; \\
P(T = t \mid X) &= \frac{\beta_P(X) + t - 2}{\alpha_P(X) + \beta_P(X) + t - 1} P(T = t - 1 \mid X); \\
P(T > 1 \mid X) &= \frac{\beta_P(X)}{\alpha_P(X) + \beta_P(X)};
\end{aligned}$$

$$P(T > t \mid X) = \frac{\beta_P(X) + t - 1}{\alpha_P(X) + \beta_P(X) + t - 1} P(T > t - 1 \mid X),$$

where $\alpha_P(X) = \exp a_P(X)$ and $\beta_P(X) = \exp b_P(X)$. Recall that our M-estimand is $\theta_P := (a_P, b_P)$, where θ_P lies in the product space $\mathcal{H} := \mathcal{H}_a \times \mathcal{H}_b$, where $\mathcal{H}_a \subseteq L^2(\mu_X)$ and $\mathcal{H}_b \subseteq L^2(\mu_X)$ for some measure μ_X on $\text{supp}(X)$. Under noninformative censoring, the log-transformed shape parameters a_P and b_P are identified via the following optimization problem:

$$(a_P, b_P) = \underset{(a,b)\in\mathcal{H}}{\text{argmin}}\, E_P[\ell(Z, a, b)],$$

where ℓ is the negative log-likelihood loss and can be written as:

$$\ell : (z, a, b) \mapsto -\delta \log P_{a,b}(T = t \mid X = x) - (1 - \delta) \log P_{a,b}(T > t \mid X = x),$$

and $P_{a,b}$ denotes the distribution of a beta-geometric random variable with shape parameters $\alpha(\cdot) = \exp a(\cdot)$ and $\beta(\cdot) = \exp b(\cdot)$.

Recall that our parameter of interest, $\Psi(P) = \psi_P(\theta_P)$, is the mean survival probability $P(T > t_0)$ at a fixed time $t_0 \in \mathbb{N}$. We can write the functional $\psi_P : \mathcal{H} \to \mathbb{R}$ as

$$\theta \mapsto \psi_P(\theta) := P_\theta(T > t_0) = E_P[m(Z, \theta)],$$

where $m : (z, \theta) \mapsto P_\theta(T > t_0 \mid X = x)$ is a nonlinear functional. The derivative of this functional at θ in the direction of $h \in \mathcal{H}$ satisfies the expression:

$$\dot{m}_\theta(h) : z \mapsto \partial_\theta P_\theta(T > t_0 \mid X = x)(h) = P_\theta(T > t_0 \mid X = x) \cdot \partial_\theta \log P_\theta(T > t_0 \mid X = x)(h),$$

where $\partial_\theta \log P_\theta(T > t_0 \mid X = x)(h)$ is determined by the recursions given later in this section.

The derivative of the functional m, as well as the gradient and Hessian operators of the loss function ℓ, can be computed using the recursion relations derived in Appendix A.2 of Hubbard et al. [2021]. For completeness, we restate these recursions here and use them to derive the relevant derivative terms. Let $\theta = (a, b) \in \mathcal{H}$ and set $\alpha := \exp a$ and $\beta := \exp b$. For a direction $(h_a, h_b) \in \mathcal{H}$, the derivative of the loss function ℓ at θ is given by:

$$\dot{\ell}_\theta(h_a, h_b)(z) := -\delta \partial_\theta \log P_\theta(T = t \mid X = x)(h_a, h_b) - (1 - \delta) \partial_\theta \log P_\theta(T > t \mid X = x)(h_a, h_b).$$

Specifically, the derivative of the log event probabilities satisfies:

$$\begin{aligned}
\partial_\theta \log P_\theta(T = 1 \mid X)(h_a, h_b) &= \frac{\beta(X)}{\alpha(X) + \beta(X)} h_a(X) - \frac{\beta(X)}{\alpha(X) + \beta(X)} h_b(X), \\
\partial_\theta \log P_\theta(T = t \mid X)(h_a, h_b) &= \partial_\theta \log P_\theta(T = t - 1 \mid X)(h_a, h_b) \\
&\quad - \frac{\alpha(X)}{\alpha(X) + \beta(X) + t - 1} h_a(X) \\
&\quad + \frac{(\alpha(X) + 1)\beta(X)}{(\beta(X) + t - 2)(\alpha(X) + \beta(X) + t - 1)} h_b(X).
\end{aligned}$$

Similarly, the derivative of the log survival probabilities satisfies:

$$\begin{aligned}
\partial_\theta \log P_\theta(T > 1 \mid X)(h_a, h_b) &= -\frac{\alpha(X)}{\alpha(X) + \beta(X)} h_a(X) + \frac{\alpha(X)}{\alpha(X) + \beta(X)} h_b(X), \\
\partial_\theta \log P_\theta(T > t \mid X)(h_a, h_b) &= \partial_\theta \log P_\theta(T > t - 1 \mid X)(h_a, h_b) \\
&\quad - \frac{\alpha(X)}{\alpha(X) + \beta(X) + t - 1} h_a(X)
\end{aligned}$$

$$+ \frac{\alpha(X)\beta(X)}{(\beta(X)+t-1)(\alpha(X)+\beta(X)+t-1)} h_b(X).$$

Furthermore, for two directions $\theta_1 : (h_a, h_b) \in \mathcal{H}$ and $\theta_2 := (g_a, g_b) \in \mathcal{H}$, the second derivative of the loss function ℓ is given by:

$$\ddot{\ell}_\theta(\theta_1, \theta_2)(z) := -\delta \partial_\theta^2 \log P_\theta(T = t \mid X = x)(\theta_1, \theta_2) - (1-\delta)\partial_\theta^2 \log P_\theta(T > t \mid X = x)(\theta_1, \theta_2).$$

To determine the second derivatives on the right-hand side, we use the fact that, for any functional $\theta \mapsto f_\theta(t, a, x)$, its second derivative satisfies the Jacobian formula:

$$\begin{aligned}\partial_\theta^2 f_\theta(t,a,x)(\theta_1,\theta_2) = {} & \partial_a^2 f_\theta(t,a,x)(h_a,g_a) + \partial_b^2 f_\theta(t,a,x)(h_b,g_b) \\ & + \partial_a\partial_b f_\theta(t,a,x)(h_a,g_b) + \partial_b\partial_a f_\theta(t,a,x)(h_b,g_a).\end{aligned}$$

Applying the chain rule and using that $\partial_a \alpha(X) = \partial_a \exp\{a(X)\} = \alpha(X)$ and $\partial_b \beta(X) = \partial_b \exp\{b(X)\} = \beta(X)$, we can show that

$$\begin{aligned}\partial_a^2 \log P_\theta(T=1 \mid X=x)(h_a,g_a) &= -\frac{\alpha(X)\beta(X)}{\{\alpha(X)+\beta(X)\}^2} h_a(X) g_a(X); \\ \partial_a\partial_b \log P_\theta(T=1 \mid X=x)(h_a,g_b) &= \frac{\alpha(X)\beta(X)}{\{\alpha(X)+\beta(X)\}^2} h_a(X) g_b(X); \\ \partial_b\partial_a \log P_\theta(T=1 \mid X=x)(h_b,g_a) &= \frac{\alpha(X)\beta(X)}{\{\alpha(X)+\beta(X)\}^2} h_b(X) g_a(X); \\ \partial_b^2 \log P_\theta(T=1 \mid X=x)(h_b,g_b) &= -\frac{\alpha(X)\beta(X)}{\{\alpha(X)+\beta(X)\}^2} h_b(X) g_b(X).\end{aligned}$$

Furthermore, applying the recursions, we find that

$$\begin{aligned}\partial_a^2 \log P_\theta(T=t \mid X=x)(h_a,g_a) &= \partial_a^2 \log P_\theta(T=t-1 \mid X=x)(h_a,g_a) \\ &\quad + \left[\frac{-\alpha(X)(\beta(X)+t-1)}{(\alpha(X)+\beta(X)+t-1)^2}\right] h_a(X) g_a(X); \\ \partial_b^2 \log P_\theta(T=t \mid X=x)(h_b,g_b) &= \partial_b^2 \log P_\theta(T=t-1 \mid X=x)(h_b,g_b) \\ &\quad + \beta(X)(\alpha(X)+1) \cdot \left[\frac{1}{(\beta(X)+t-2)(\alpha(X)+\beta(X)+t-1)}\right. \\ &\quad \left. - \frac{\beta(X)(2\beta(X)+\alpha(X)+2t-3)}{[(\beta(X)+t-2)(\alpha(X)+\beta(X)+t-1)]^2}\right] h_b(X) g_b(X) \\ \partial_a\partial_b \log P_\theta(T=t \mid X=x)(h_a,g_b) &= \partial_a\partial_b \log P_\theta(T=t-1 \mid X=x)(h_a,g_b) \\ &\quad + \frac{\alpha(X)\beta(X)}{\{\alpha(X)+\beta(X)+t-1\}^2} h_a(X) g_b(X) \\ \partial_b\partial_a \log P_\theta(T=t \mid X=x)(h_b,g_a) &= \partial_b\partial_a \log P_\theta(T=t-1 \mid X=x)(h_b,g_a) \\ &\quad + \frac{\alpha(X)\beta(X)}{\{\alpha(X)+\beta(X)+t-1\}^2} h_b(X) g_a(X)\end{aligned}$$

Similarly, we can compute the second derivatives of the log survival probabilities. At $t = 1$, they satisfy

$$\begin{aligned}\partial_a^2 \log P_\theta(T>1 \mid X=x)(h_a,g_a) &= -\frac{\alpha(X)\beta(X)}{\{\alpha(X)+\beta(X)\}^2} h_a(X) g_a(X); \\ \partial_a\partial_b \log P_\theta(T>1 \mid X=x)(h_a,g_b) &= \frac{\alpha(X)\beta(X)}{\{\alpha(X)+\beta(X)\}^2} h_a(X) g_b(X); \\ \partial_b\partial_a \log P_\theta(T>1 \mid X=x)(h_b,g_a) &= \frac{\alpha(X)\beta(X)}{\{\alpha(X)+\beta(X)\}^2} h_b(X) g_a(X);\end{aligned}$$

$$\partial_b^2 \log P_\theta(T > 1 \mid X = x)(h_b, g_b) = -\frac{\alpha(X)\beta(X)}{\{\alpha(X) + \beta(X)\}^2} h_b(X) g_b(X).$$

Moreover, using the identity

$$\log P_\theta(T = t \mid X = x) = a(X) - \log\{\beta(X) + t - 1\} + \log P_\theta(T > t \mid X = x),$$

we obtain, for $t > 1$,

$$\begin{aligned}
\partial_a^2 \log P_\theta(T > t \mid X = x)(h_a, g_a) &= \partial_a^2 \log P_\theta(T = t \mid X = x)(h_a, g_a), \\
\partial_a \partial_b \log P_\theta(T > t \mid X = x)(h_a, g_b) &= \partial_a \partial_b \log P_\theta(T = t \mid X = x)(h_a, g_b), \\
\partial_b \partial_a \log P_\theta(T > t \mid X = x)(h_b, g_a) &= \partial_b \partial_a \log P_\theta(T = t \mid X = x)(h_b, g_a), \\
\partial_b^2 \log P_\theta(T > t \mid X = x)(h_b, g_b) &= \partial_b^2 \log P_\theta(T = t \mid X = x)(h_b, g_b) \\
&\quad + \frac{\beta(X)(t-1)}{\{\beta(X) + t - 1\}^2} h_b(X) g_b(X).
\end{aligned}$$

4.F.3 Additional details on synthetic experiment

The data structure $Z = (X, A, \widetilde{T}, \Delta)$ is generated as follows. The covariate vector X is drawn from a uniform distribution on $[-1, 1]^3$. Given X, the treatment assignment A is sampled from a Bernoulli distribution with success probability

$$\pi_0(X) := \text{expit}(2\sqrt{3}(X_1 + X_2 + X_3)).$$

The uncensored survival time T is drawn conditional on (A, X) from a Beta-geometric event time distribution with shape parameters

$$\begin{aligned}
\alpha_0(A, X) &:= \exp(-0.1 + \sqrt{3}(X_1 + X_2 + X_3) + 0.25A) \\
\beta_0(A, X) &:= \exp(\sqrt{3}(X_1 + X_2 + X_3) + 0.1).
\end{aligned}$$

We impose administrative censoring at time $c = 6$, with the observed event time $\widetilde{T} = \min(T, c)$ and censoring indicator $\Delta = 1(T \leq c)$.

4.G Additional experiments for R-learner loss

4.G.1 Inference on ATE using R-learner

In this experiment, we evaluate various autoDML estimators for the ATE derived from the R-learner loss, a specific Neyman-orthogonal loss for the CATE introduced by Nie and Wager [2021]. One advantage of developing autoDML estimators based on a loss function that directly targets the CATE is that it provides a straightforward framework for constructing semiparametric estimators with model assumptions on the CATE, such as partially linear regression models, by selecting an appropriate Hilbert space $\mathcal{H}$. In this experiment, we impose an additive model assumption on the CATE by defining $\mathcal{H}$ as the space of all functions that are additive in the covariates.

We consider the data structure $Z = (X, A, Y)$, where $X \in \mathbb{R}^d$ represents covariates, $A \in \{0, 1\}$ is a binary treatment, and $Y \in \mathbb{R}$ is the outcome. The R-learner loss function is a weighted pseudo-regression loss, given in Example 1c. The M-estimand θ_P is the CATE function, $x \mapsto E_P[Y \mid A = 1, X = x] - E_P[Y \mid A = 0, X = x]$, with the functional of interest $\psi_P : \theta \mapsto E_P[\theta(X)]$. The covariate vector $X = (X_1, X_2, X_3)$ is drawn independently

as $X_1, X_2, X_3 \sim \text{Uniform}(-1, 1)$. Given $(X_1, X_2, X_3) = (x_1, x_2, x_3)$, treatment A follows a Bernoulli distribution with success probability $\pi_0(x_1, x_2, x_3) := \text{expit}(\sin(2x_1) + 2x_2 + |x_3|)$. We define the auxiliary probability $\pi_{0,\text{add}}(x_2) := \text{expit}(x_2)$ to construct a local perturbation, where the weight is $\omega(x_2) = \pi_{0,\text{add}}(x_2)(1 - \pi_{0,\text{add}}(x_2))$ and the local fluctuation term is $\text{local\_fluct} = (A - \pi_{0,\text{add}}(X_2))/\omega(X_2)$. The outcome Y, given $A = a$ and $(X_1, X_2, X_3) = (x_1, x_2, x_3)$, follows a Normal distribution with mean $\mu_0(a, x_1, x_2, x_3) = \text{local\_fluct}/\sqrt{n} + \sin(2x_1) + |x_3| + a(0.3 + 2x_3^2 + \sin(2x_1))$ and standard deviation $\sigma_0(x_1, x_3) = 0.5 + (x_1 + x_3)/8$. Irregular estimators exhibit nonvanishing asymptotic bias under sampling from local perturbations of P_0 at distance $O(n^{-1/2})$.

We evaluate four autoDML estimators based on the R-learner loss, each employing an additive model $\mathcal{H}$ for the CATE. For all estimators, the nuisance functions for the R-learner loss, specifically the propensity score π_0 and the conditional mean of the outcome given covariates $E_0[Y \mid X = \cdot]$, are estimated using an ensemble of multivariate adaptive regression splines [Friedman, 1991] and a generalized additive spline model [Hastie and Tibshirani, 1987], respectively implemented in the `R` packages `earth` [Milborrow, 2019] and `mgcv` [Wood, 2001]. The first estimator is the one-step autoDML estimator, $\widehat{\psi}_n^{\text{dml}}$, as described in Section 4.4.1 and implemented according to Algorithm 8. The second is the autoTMLE estimator, $\widehat{\psi}_n^{\text{tmle}}$, which is also based on the R-learner loss and implemented according to Algorithm 9. For both the autoDML and autoTMLE estimators, the CATE, θ_0, is learned using the R-learner loss with the highly adaptive lasso (HAL) spline estimator [van der Laan, 2015, Benkeser and van der Laan, 2016], while the Riesz representer, α_0, is learned using the Riesz loss with the HAL estimator. The HAL estimators are implemented using the `hal9001` [Hejazi et al., 2020] package, with the additive model constraint enforced by specifying the argument `max_degree = 1`. The estimators of the nuisances in the loss, the CATE, and the Riesz representer are cross-fitted using Algorithm 7 with 10 folds. The third estimator, $\widehat{\psi}_n^{\text{sieve}}$, is an autoSieve estimator that leverages a data-adaptive sieve, $\{\mathcal{H}_{n,\lambda_k} : k \in \mathbb{N}\}$, learned using the additive HAL estimator and trained without cross-fitting on all available data. Here, $\lambda_k \in \mathbb{R}$ represents the lasso regularization parameter, and $\mathcal{H}_{n,\lambda_k}$ denotes the linear span of basis functions with nonzero coefficients from the additive HAL estimator fit with regularization parameter λ_k. The regularization parameter $\lambda_{k(n)}$ used to compute $\widehat{\psi}_n^{\text{sieve}}$ is selected data-adaptively using the automatic undersmoothing method detailed in Section 4.4.1. The final estimator is an adaptive sieve plug-in estimator for the ATE, where the sieve is again learned using HAL but without undersmoothing, resulting in superefficient behavior. This estimator is a specific instance of the adaptive DML estimators introduced in van der Laan et al. [2023] and is also a special case of the estimators discussed in Appendix 4.E, corresponding to a HAL-based sieve estimator fit with the regularization parameter $\lambda_{k(n)}$ selected through cross-validation using the R-learner loss. As a baseline for comparison, we also evaluate the nonparametric AIPW estimator of the ATE, which is constructed using estimators of the propensity score and the outcome regression.

The results are displayed in Figure 4.3. In terms of bias and coverage, autoTMLE, autoDML, and DML perform best among the methods evaluated. autoTMLE and autoDML exhibit less variability compared to DML because they perform debiasing within a semiparametric model and are thus less sensitive to issues with treatment overlap. Additionally, autoDML and autoTMLE demonstrate similar performance across all sample sizes, likely

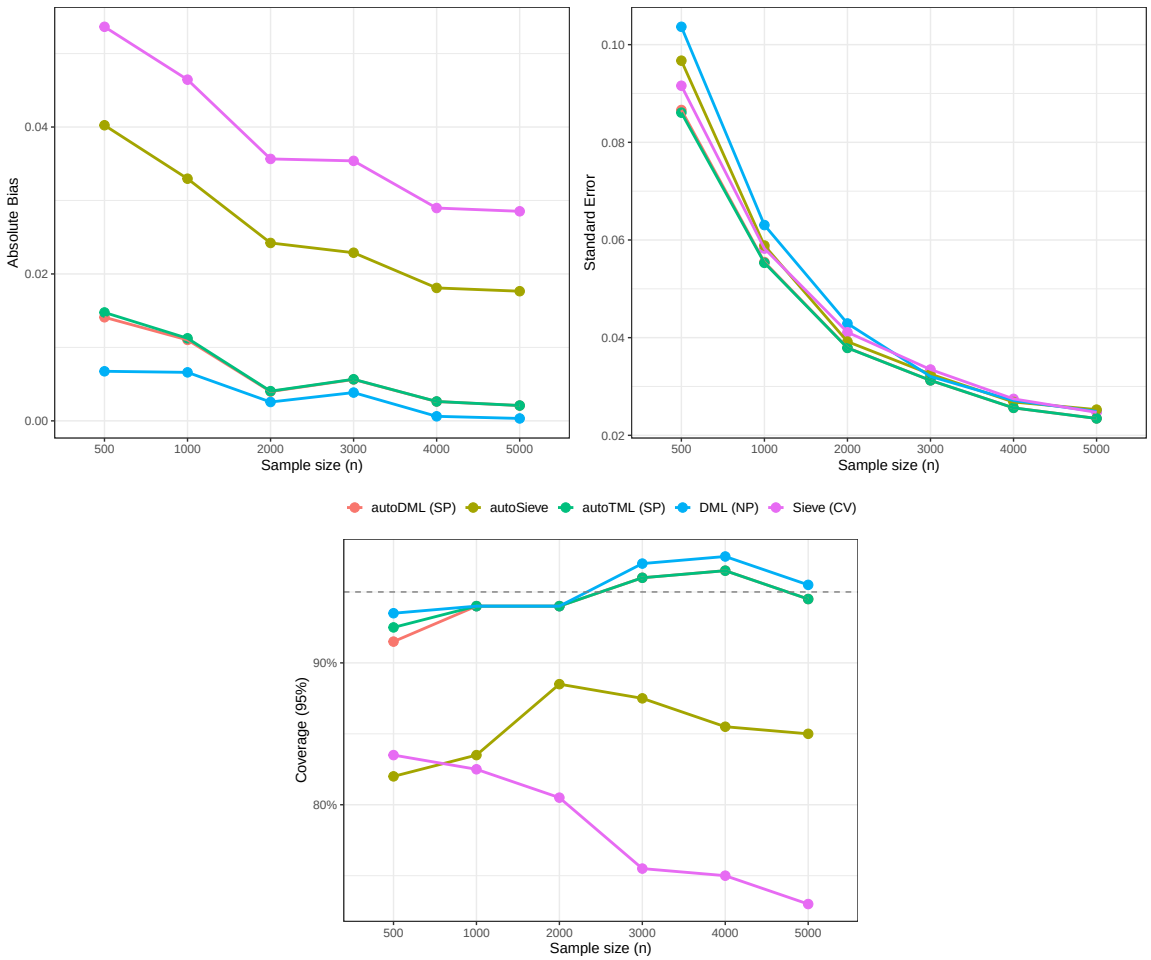


Figure 4.3: Bias, standard error, and coverage probability for the ATE, comparing various methods across sample sizes.

due to the quadratic nature of the loss and the linearity of the ATE functional, which aligns the two approaches closely. Conversely, the cross-validated and auto-undersmoothed plug-in HAL estimators exhibit the poorest performance regarding bias and coverage. The cross-validated estimator shows significant bias and inadequate coverage under sampling from local perturbations, primarily due to its irregular nature. Although undersmoothing partially reduces this bias, it remains substantial enough to negatively affect coverage.

Chapter 5

CONCLUSION

Appendix A

PROOFS FOR CHAPTER 1

A.1 Proofs for Additional discussion of conditions

Proof. Write

$$E_P[f(Z)] - E_{P'}[f(Z)] = \int f(z)\{p(z) - p'(z)\}\, d\nu(z) = \int f(z)\big(\sqrt{p}(z) - \sqrt{p'(z)}\big)\big(\sqrt{p}(z) + \sqrt{p'(z)}\big)\, d\nu(z).$$

By Cauchy–Schwarz,

$$\big|E_P[f(Z)] - E_{P'}[f(Z)]\big| \le \left(\int f(z)^2\big(\sqrt{p}(z) + \sqrt{p'(z)}\big)^2\, d\nu(z)\right)^{1/2} \left(\int \big(\sqrt{p}(z) - \sqrt{p'(z)}\big)^2\, d\nu(z)\right)^{1/2}.$$

For the first factor, use $2\sqrt{pp'} \le p + p'$ to obtain

$$\int f^2(\sqrt{p} + \sqrt{p'})^2\, d\nu = \int f^2\{p + p' + 2\sqrt{pp'}\}\, d\nu \le 2\int f^2(p + p')\, d\nu = 2\{E_P[f(Z)^2] + E_{P'}[f(Z)^2]\}.$$

For the second factor,

$$\left(\int (\sqrt{p} - \sqrt{p'})^2\, d\nu\right)^{1/2} = \sqrt{2}\, H(P, P').$$

Combining these bounds yields

$$\big|E_P[f(Z)] - E_{P'}[f(Z)]\big| \le \big(2\{E_P[f^2] + E_{P'}[f^2]\}\big)^{1/2} \cdot \sqrt{2}\, H(P, P') = 2\, H(P, P')\left(E_P[f^2] + E_{P'}[f^2]\right)^{1/2},$$

as claimed. □

A.2 Proofs for Sections 2.2.2, 2.3.2, and 2.4.1

A.2.1 Proof of Theorem 2.1

Proof of Theorem 2.1. We will first show that $(P, \mu) \mapsto \psi_P(\mu)$ is total pathwise differentiable in the sense of Appendix A.4. in Luedtke [2024]. To do so, we apply Lemma S6 in Luedtke [2024]. To do so, we need to show that the total operator $(P, \mu) \mapsto \psi_P(\mu)$ is partially differentiable in its first argument and continuously partially differentiable in its second argument at (P_0, μ_0), where these notions of differentiability are defined in Appendix A.4 of Luedtke [2024]. The former differentiability holds since, by assumption A3, $P \mapsto \psi_P(\mu_0)$ is pathwise differentiable at P_0. To establish continuous partial differentiability in the second argument, note that $\mu \mapsto \psi_P(\mu)$ is a linear functional and, thus, the partial derivative of $(P, \mu) \mapsto \psi_P(\mu)$ in its second argument is simply $\mu \mapsto \psi_P(\mu)$. Hence, to establish continuous partial differentiability, it remains to show that $P \mapsto \psi_P$ is a continuous functional at P_0 with the domain equipped with the Hellinger distance and the range equipped with the operator norm. This is true by assumption A2. By Lemma S6 in Luedtke [2024], the total pathwise derivative at (P_0, μ_0) in the direction of $(s, h) \in T_{\mathcal{M}}(P_0) \times \mathcal{H}$ is then given by

$\langle \phi_{0,\mu_0}, s\rangle + \langle \alpha_0, h\rangle$.

Under our assumptions, $P \mapsto \mu_P$ is pathwise differentiable at P_0 in a Hilbert-valued sense by Example 5 of Luedtke and Chung [2024], with pathwise derivative $d\mu_0(s)$ in the direction $s \in T_{\mathcal{M}}(P_0)$ given by $h \mapsto \langle h\{\mathcal{I}_Y - \mu_0\}, s\rangle$, where $\mathcal{I}_Y : o \mapsto y$. As a consequence, by the chain rule, $P \mapsto \psi_P(\mu_P)$ is pathwise differentiable at P_0 with pathwise derivative given by $\langle \phi_{0,\mu_0}, s\rangle + \langle \alpha_0\{\mathcal{I}_Y - \mu_0\}, s\rangle$. It follows that the efficient influence function is $\phi_{0,\mu_0} + D_{\mu_0,\alpha_0}$. □

A.2.2 Proofs for Theorem 2.3 and Lemma 2.4

Proof of Theorem 2.3. We begin with the proof of the Riesz-based decomposition, as the regression-based decomposition will follow as a special case.

By the law of iterated expectations, conditioning on $(\alpha_n(A,W), \alpha_0(A,W))$ under P_0 yields

$$\langle \alpha_0 - \alpha_n, \mu_n - \mu_0\rangle = \big\langle \alpha_0 - \alpha_n,\ \Pi_{(\alpha_n,\alpha_0)}(\mu_n - \mu_0)\big\rangle.$$

We decompose the preceding display as

$$\begin{aligned}\langle \alpha_0 - \alpha_n, \mu_n - \mu_0\rangle &= \big\langle \alpha_0 - \alpha_n,\ \Pi_{\alpha_n}(\mu_n - \mu_0)\big\rangle + \big\langle (\Pi_{(\alpha_n,\alpha_0)} - \Pi_{\alpha_n})(\alpha_0 - \alpha_n),\ (\mu_n - \mu_0) - \Pi_{\alpha_n}(\mu_n - \mu_0)\big\rangle \\ &= \big\langle \alpha_n - \alpha_0,\ r_{n,0}\big\rangle - \big\langle (\Pi_{(\alpha_n,\alpha_0)} - \Pi_{\alpha_n})\Delta_{\alpha,n},\ \Delta_{\mu,n} - \Pi_{\alpha_n}\Delta_{\mu,n}\big\rangle,\end{aligned} \tag{A.1}$$

where $r_{n,0} = -\Pi_{\alpha_n}(\mu_n - \mu_0)$. By the Riesz representation property for α_0 in (2.1), the first term on the right-hand side can be expressed as

$$\begin{aligned}\big\langle \alpha_n - \alpha_0,\ r_{n,0}\big\rangle &= -\psi_0(r_{n,0}) + P_0\{\alpha_n r_{n,0}\} \\ &= -\psi_n(r_{n,0}) + P_0\{\alpha_n r_{n,0}\} - \psi_0(r_{n,0}) + \psi_n(r_{n,0}),\end{aligned}$$

where, in the second equality, we add and subtract $\psi_n(r_{n,0})$. By the orthogonality condition in (2.5),

$$\psi_n(r_{n,0}) = P_n\{\alpha_n r_{n,0}\}.$$

Hence,

$$-\psi_n(r_{n,0}) + P_0\{\alpha_n r_{n,0}\} = -(P_n - P_0)\{\alpha_n r_{n,0}\}.$$

Therefore,

$$\begin{aligned}\big\langle \alpha_n - \alpha_0,\ r_{n,0}\big\rangle &= (P_n - P_0)\{-\alpha_n r_{n,0}\} - \psi_0(r_{n,0}) + \psi_n(r_{n,0}) \\ &= (P_n - P_0)\{-\alpha_n r_{n,0} + \phi_{r_{n,0}}\} + \psi_n(r_{n,0}) - \psi_0(r_{n,0}) - P_n\phi_{r_{n,0}} \\ &= (P_n - P_0)B_{n,0} + Rem_{\psi_n}(r_{n,0}),\end{aligned}$$

where we add and subtract $P_n\phi_{r_{n,0}}$ and use that $P_0\phi_{r_{n,0}} = 0$. Substituting this display into (A.1), we obtain

$$\begin{aligned}\langle \alpha_0 - \alpha_n, \mu_n - \mu_0\rangle &= (P_n - P_0)B_{n,0} + Rem_{\psi_n}(r_{n,0}) - \big\langle (\Pi_{(\alpha_n,\alpha_0)} - \Pi_{\alpha_n})\Delta_{\alpha,n},\ \Delta_{\mu,n} - \Pi_{\alpha_n}\Delta_{\mu,n}\big\rangle \\ &= (P_n - P_0)B_{n,0} + \big\langle \alpha_0 - \Pi_{\alpha_n}\alpha_0,\ \Delta_{\mu,n} - \Pi_{\alpha_n}\Delta_{\mu,n}\big\rangle + Rem_{\psi_n}(r_{n,0}),\end{aligned} \tag{A.2}$$

where we use that

$$(\Pi_{(\alpha_n,\alpha_0)} - \Pi_{\alpha_n})\Delta_{\alpha,n} = \Pi_{\alpha_n}\alpha_0 - \alpha_0 = -(\alpha_0 - \Pi_{\alpha_n}\alpha_0).$$

The Riesz-based decomposition now follows. If Condition B1ii holds, then

$$\big\langle (\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})\Delta_{\alpha,n},\ \mu_0 - \Pi_{\mu_n}\mu_0\big\rangle = \gamma_{\alpha,0}(\mu_n,\alpha_n)\,\|\alpha_n - \alpha_0\|^2 = O_p\left(\|\alpha_n - \alpha_0\|^2\right).$$

By the Cauchy–Schwarz inequality, we also have

$$\left\langle (\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})\Delta_{\alpha,n},\ \mu_0 - \Pi_{\mu_n}\mu_0 \right\rangle \leq \|(\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})\Delta_{\alpha,n}\|\, \|\mu_0 - \Pi_{\mu_n}\mu_0\|.$$

By nonexpansiveness of orthogonal projections, $\|(\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})\Delta_{\alpha,n}\| \leq \|\Delta_{\alpha,n}\|$ and

$$\|\mu_0 - \Pi_{\mu_n}\mu_0\| = \|\Delta_{\mu,n} - \Pi_{\mu_n}\Delta_{\mu,n}\| \leq \|\Delta_{\mu,n}\|.$$

Hence,

$$\left\langle (\Pi_{(\mu_n,\mu_0)} - \Pi_{\mu_n})\Delta_{\alpha,n},\ \mu_0 - \Pi_{\mu_n}\mu_0 \right\rangle \leq \|\Delta_{\alpha,n}\|\, \|\Delta_{\mu,n}\|.$$

Taking the minimum of these two valid bounds, we obtain

$$\left\langle (\Pi_{(\alpha_n,\alpha_0)} - \Pi_{\alpha_n})\Delta_{\mu,n},\ \Delta_{\alpha,n} \right\rangle = O_p\left(\|\alpha_n - \alpha_0\|^2 \wedge \|\mu_n - \mu_0\|\, \|\Pi_{\alpha_n}\alpha_0 - \alpha_0\|\right),$$

as desired.

For the regression-based decomposition, we apply the Riesz-based decomposition with $\psi_0(\alpha) := E_0\{\alpha(A,W)\mu_0(A,W)\}$ and $\psi_n(\alpha) := \frac{1}{n}\sum_{i=1}^n \alpha(A_i,W_i)Y_i$ (see also the proof outline in Section 2.3.2). In particular, μ_0 is the Riesz representer of the linear functional $\psi_0 : \alpha \mapsto E_0\{\alpha(A,W)\mu_0(A,W)\}$. Moreover, the orthogonality condition in (2.4) is equivalent to the orthogonality condition (2.5) for the estimated linear functional ψ_n. For this choice of ψ_n, we have $Rem_{\psi_n}(\alpha) = 0$ with $\phi_\alpha(z) := \alpha y - \psi_0(\alpha)$. Under these definitions, B_0 and $B_{n,0}$ coincide with A_0 and $A_{n,0}$, respectively. The regression-based decomposition now follows.

□

Proof of Lemma 2.4. The proof that Conditions B1i and B1ii are implied by B1*i and B1*ii, respectively, follows directly from the Cauchy–Schwarz inequality applied to the covariances in the numerators of $\gamma_{\mu,0}(\mu_n,\alpha_n)$ and $\gamma_{\alpha,0}(\mu_n,\alpha_n)$. It therefore remains to show that B1*i and B1*ii are implied by B1**i and B1**ii, respectively. We prove that B1**i implies B1*i; the remaining implication follows by symmetry.

Denote $X := (W,A)$. Let $g : \mathbb{R}^2 \to \mathbb{R}$ be a Lipschitz continuous function with constant $L > 0$. By Lipschitz continuity, we have that

$$\begin{aligned}
&|g(\mu_n(x),\mu_0(x)) - E[g(\mu_n(X),\mu_0(X))|\mu_n(X) = \mu_n(x)]| \\
&\qquad = |E[g(\mu_n(x),\mu_0(x)) - g(\mu_n(x),\mu_0(X))|\mu_n(X) = \mu_n(x)]| \\
&\qquad \leq E[|g(\mu_n(x),\mu_0(x)) - g(\mu_n(x),\mu_0(X))|\,|\mu_n(X) = \mu_n(x)] \\
&\qquad \leq LE[|\mu_0(x) - \mu_0(X)|\,|\mu_n(X) = \mu_n(x)].
\end{aligned}$$

On the event $\{\mu_n(X) = \mu_n(x)\}$, we know

$$\begin{aligned}
|\mu_0(x) - \mu_0(X)| &\leq |\mu_0(x) - \mu_n(x)| + |\mu_n(x) - \mu_n(X)| + |\mu_0(X) - \mu_n(X)| \\
&\leq |\mu_0(x) - \mu_n(x)| + |\mu_0(X) - \mu_n(X)|.
\end{aligned}$$

Therefore, it holds that

$$\begin{aligned}
&|g(\mu_n(x),\mu_0(x)) - E[g(\mu_n(X),\mu_0(X))|\mu_n(X) = \mu_n(x)]| \\
&\qquad \lesssim E[|\mu_0(x) - \mu_0(X)|\ |\ \mu_n(X) = \mu_n(x)] \\
&\qquad \lesssim E|\mu_0(x) - \mu_n(x)| + E[|\mu_0(X) - \mu_n(X)|\ |\ \mu_n(X) = \mu_n(x)].
\end{aligned}$$

Now, for some function $f \in L^2(P_X)$, suppose that $(\mu_n(x),\mu_0(x)) \mapsto (\Pi_{(\mu_n,\mu_0)}f)(x)$ is Lipschitz continuous. Then, defining $g : (\widehat{m},m) \mapsto E_0[f(X)|\mu_n(X) = \widehat{m}, \mu_0(X) = m, \mathcal{D}_n]$

and noting by the law of iterated expectation that $\Pi_{\mu_n}\Pi_{(\mu_n,\mu_0)}f = \Pi_{\mu_n}f$, we obtain the following pointwise error bound:

$$\left|\Pi_{(\mu_n,\mu_0)}f - \Pi_{\mu_n}f\right| \lesssim |\mu_n - \mu_0| + \Pi_{\mu_n}(|\mu_n - \mu_0|).$$

Since $\|\Pi_{\mu_n}(|\mu_n - \mu_0|)\| \leq \|\mu_n - \mu_0\|$ by the properties of projections, it follows that

$$\|\Pi_{(\mu_n,\mu_0)}f - \Pi_{\mu_n}f\| \lesssim \|\mu_n - \mu_0\|.$$

Taking f to be $\Delta_{\alpha,n}$, we conclude that Condition B1**i implies Condition B1*i. By an identical argument, Condition B1**ii implies Condition B1*ii. The result then follows.

By symmetry of μ_n and μ_0 in the Lipschitz continuity condition and the preceding bound, we also have

$$\|\Pi_{(\mu_n,\mu_0)}f - \Pi_{\mu_0}f\| \lesssim \|\mu_n - \mu_0\|.$$

Hence, by the triangle inequality,

$$\begin{aligned}\|\Pi_{\mu_0}f - \Pi_{\mu_n}f\| &\leq \|\Pi_{(\mu_n,\mu_0)}f - \Pi_{\mu_0}f\| + \|\Pi_{(\mu_n,\mu_0)}f - \Pi_{\mu_n}f\| \\ &\lesssim \|\mu_n - \mu_0\|.\end{aligned}$$

□

A.2.3 Proof of Lemma 2.5

Proof of Lemma 2.5. The orthogonality conditions follow directly from the first-order optimality conditions of (2.10) and (2.11) for the least-squares and Riesz losses. For the least-squares loss, note that $\mu_n = \theta_n \circ \mu_n$, where

$$\theta_n \in \operatorname*{argmin}_{\theta} \sum_{i=1}^{n} \{Y_i - \theta(\mu_n(A_i, W_i))\}^2.$$

By the first-order optimality conditions for θ_n, for each $\theta : \mathbb{R} \to \mathbb{R}$,

$$\begin{aligned}0 &= \frac{d}{d\varepsilon} \sum_{i=1}^{n} \Big\{Y_i - (\theta_n + \varepsilon\theta)\big(\mu_n(A_i, W_i)\big)\Big\}^2 \Big|_{\varepsilon=0} \\ &= 2\,\frac{1}{n} \sum_{i=1}^{n} \theta\big(\mu_n(A_i, W_i)\big)\{Y_i - \mu_n(A_i, W_i)\}.\end{aligned}$$

Hence, (2.12) holds, and choosing θ such that $\theta \circ \mu_n = r_{n,0}$, we conclude that μ_n satisfies (2.4).

The implication for the Riesz loss follows by an identical argument. Here we use that, by B2, ψ_n is a bounded linear functional on $L^2(P_n)$, and

$$\begin{aligned}\frac{d}{d\varepsilon}\psi_n\left(Z_i, \alpha_n + \varepsilon\,\theta \circ \alpha_n\right)\Big|_{\varepsilon=0} &= \frac{d}{d\varepsilon}\psi_n\left(Z_i, \varepsilon\,\theta \circ \alpha_n\right)\Big|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon}\,\varepsilon\,\psi_n\left(Z_i, \theta \circ \alpha_n\right)\Big|_{\varepsilon=0} \\ &= \psi_n\left(Z_i, \theta \circ \alpha_n\right).\end{aligned}$$

See also Lemma A.16, which establishes empirical calibration for the isocalibrated nuisance estimators. □

A.2.4 Proofs of Lemma 2.10 and Lemma 2.11

Proof of Lemma 2.10. Let $\mathcal{I}_Y : z \mapsto y$ be the coordinate projection onto the outcome variable. By the triangle inequality, we have

$$\begin{aligned}\|A_{n,0} - A_0\| \leq \|&\Pi_{\mu_n}(\alpha_0 - \overline{\alpha}_0)\{\mathcal{I}_Y - \mu_0(z)\} - \Pi_{\mu_0}(\alpha_0 - \overline{\alpha}_0)\{\mathcal{I}_Y - \mu_0(z)\}\| \\ &+ \|\Pi_{\mu_n}(\overline{\alpha}_0 - \alpha_n)\{\mathcal{I}_Y - \mu_n(z)\} - \Pi_{\mu_n}(\alpha_0 - \overline{\alpha}_0)\{\mathcal{I}_Y - \mu_0(z)\}\|.\end{aligned}$$

By boundedness of nuisances in G1, consistency of μ_n for μ_0, and the projection bound in G2, the first term satisfies

$$\begin{aligned}\|\Pi_{\mu_n}(\alpha_0 - \overline{\alpha}_0)\{\mathcal{I}_Y - \mu_0(z)\} - \Pi_{\mu_0}(\alpha_0 - \overline{\alpha}_0)\{\mathcal{I}_Y - \mu_0(z)\}\| &= \|\{(\Pi_{\mu_n} - \Pi_{\mu_0})(\alpha_0 - \overline{\alpha}_0)\}\{\mathcal{I}_Y - \mu_0(z)\}\| \\ &= O_p(\|(\Pi_{\mu_n} - \Pi_{\mu_0})(\alpha_0 - \overline{\alpha}_0)\|) \\ &= O_p(\|\mu_n - \mu_0\|^{\beta}) \\ &= o_p(1).\end{aligned}$$

By boundedness of nuisances in G1 and nonexpansiveness of projections, the second term satisfies:

$$\begin{aligned}&\|\Pi_{\mu_n}(\overline{\alpha}_0 - \alpha_n)\{\mathcal{I}_Y - \mu_n(z)\} - \Pi_{\mu_n}(\alpha_0 - \overline{\alpha}_0)\{\mathcal{I}_Y - \mu_0(z)\}\|. \\ &\quad\lesssim O_p(\|\mu_n - \mu_0\|) + \|\{\Pi_{\mu_n}(\overline{\alpha}_0 - \alpha_n)\}\{\mathcal{I}_Y - \mu_0(z)\}\| \\ &\quad\lesssim O_p(\|\mu_n - \mu_0\|) + O_p(\|\overline{\alpha}_0 - \alpha_n\|).\end{aligned}$$

Thus, by consistency of μ_n for μ_0 and convergence of α_n to the misspecified limit $\overline{\alpha}_0$ in G2,

$$\|\Pi_{\mu_n}(\overline{\alpha}_0 - \alpha_n)\{\mathcal{I}_Y - \mu_n(z)\} - \Pi_{\mu_n}(\alpha_0 - \overline{\alpha}_0)\{\mathcal{I}_Y - \mu_0(z)\}\| = o_p(1).$$

We conclude that

$$\|A_{n,0} - A_0\| = o_p(1).$$

By G3, $A_{n,0} - A_0$ belongs, with probability tending to one, to a fixed Donsker class. Hence, by asymptotic equicontinuity of the empirical process on Donsker classes [Van Der Vaart and Wellner, 1996] and the fact that $\|A_{n,0} - A_0\| = o_p(1)$, we have $(P_n - P_0)(A_{n,0} - A_0) = o_p(n^{-1/2})$.

□

Proof of Lemma 2.11. By the triangle inequality and boundedness in G1,

$$\begin{aligned}\|\phi_{r_{n,0}} - \phi_{r_0} - r_{n,0}\alpha_n + r_0\alpha_0\| &\lesssim \|\phi_{r_{n,0}} - \phi_{r_0}\| + \|r_{n,0}\alpha_n - r_0\alpha_0\| \\ &\lesssim \|\phi_{r_{n,0}} - \phi_{r_0}\| + \|r_{n,0} - r_0\| + \|\alpha_n - \alpha_0\|.\end{aligned}$$

By G5, $\|\phi_{r_{n,0}} - \phi_{r_0}\| = O_p(\|r_{n,0} - r_0\|)$. Hence,

$$\|\phi_{r_{n,0}} - \phi_{r_0} - r_{n,0}\alpha_n + r_0\alpha_0\| = O_p(\|r_{n,0} - r_0\|) + O_p(\|\alpha_n - \alpha_0\|).$$

By symmetry, applying G4 and arguing exactly as in the proof of Lemma 2.10, we have $\|r_{n,0} - r_0\| = o_p(1)$ and therefore $\|\phi_{r_{n,0}} - \phi_{r_0} - r_{n,0}\alpha_n + r_0\alpha_0\| = o_p(1)$. By G6, $\phi_{r_{n,0}} - \phi_{r_0} - r_{n,0}\alpha_n + r_0\alpha_0$ belongs, with probability tending to one, to a fixed Donsker class. Hence, by asymptotic equicontinuity of the empirical process on Donsker classes [Van Der Vaart and Wellner, 1996] and the convergence established above, we have $(P_n - P_0)(B_{n,0} - B_0) = o_p(n^{-1/2})$.

□

A.3 Proofs of supporting theory and notation

In this section, we introduce notation and collect technical lemmas used to prove Theorems 2.4 and 2.5.

A.3.1 Notation

Cross-averaging and cross-fitting notation. Let $J \in \mathbb{N}$ denote a fixed number of cross-fitting splits. Let $\mathcal{D}_n^1, \mathcal{D}_n^2, \ldots, \mathcal{D}_n^J$ be a partition of the available data $\mathcal{D}_n$ into J subsets of approximately equal size, with corresponding index sets $\mathcal{I}_n^1, \mathcal{I}_n^2, \ldots, \mathcal{I}_n^J$. For each $i \in [n]$, let $j(i) \in [J]$ denote the index of the fold containing observation i, so that $i \in \mathcal{I}_n^{j(i)}$. For a collection of fold-specific functions $\nu_{n,\diamond} := \{\nu_{n,j} : j \in [J]\}$, define the fold-averaged operators

$$\overline{P}_0 \nu_{n,\diamond} := \frac{1}{J} \sum_{j=1}^{J} P_0 \nu_{n,j}, \qquad \overline{P}_n \nu_{n,\diamond} := \frac{1}{J} \sum_{j=1}^{J} P_{n,j} \nu_{n,j}.$$

This notation allows us to view the collections of calibrated cross-fitted estimators $\alpha^*_{n,\diamond} := \{\alpha^*_{n,j} : j \in [J]\}$ and $\mu^*_{n,\diamond} := \{\mu^*_{n,j} : j \in [J]\}$ as extended functions on $\mathcal{A} \times \mathcal{W} \times [J]$. Specifically, let $j_{\mathrm{rv}} \in [J]$ be uniformly distributed and independent of the data $\mathcal{D}_n$. Then $\widetilde{Z} := (Z, j_{\mathrm{rv}}) \sim \overline{P}_0$, since $Z \sim P_0$ and $P(j_{\mathrm{rv}} = j) = 1/J$ for $j \in [J]$.

For a collection of fold-specific functions $f_\diamond := \{f_j : j \in [J]\}$ with each $f_j \in L^2(P_0)$, define the associated $L^2(\overline{P}_0)$ norm by

$$\|f_\diamond\|_{\overline{P}_0} := \{\overline{P}_0(f_\diamond^2)\}^{1/2}.$$

Under this notation, we may view $f_\diamond$ as a single function on $\mathcal{Z} \times [J]$ and hence as an element of $L^2(\overline{P}_0)$. We define the cross-averaged functionals

$$\overline{\psi}_n(f_\diamond) := \frac{1}{J} \sum_{j=1}^{J} \psi_{n,j}(f_j), \qquad \overline{\psi}_0(f_\diamond) := \frac{1}{J} \sum_{j=1}^{J} \psi_0(f_j).$$

Let $v_\diamond := \{v_j : \mathcal{A} \times \mathcal{W} \to \mathbb{R}^k\}_{j=1}^J$ be feature maps, and define the fold-averaged projection operator $\overline{\Pi}_{v_\diamond} : \mathcal{H}^J \to \mathcal{H}^J$ by

$$(\overline{\Pi}_{v_\diamond} f_\diamond)_j := \theta^\star \circ v_j, \qquad \theta^\star \in \arg\min_\theta \|f_\diamond - \theta \circ v_\diamond\|^2_{\overline{P}_0},$$

where the minimization is over measurable $\theta : \mathbb{R}^k \to \mathbb{R}$ such that $\|\theta \circ v_\diamond\|_{\overline{P}_0} < \infty$. Thus, $\overline{\Pi}_{v_\diamond} f_\diamond$ is the $L^2(\overline{P}_0)$ projection of $f_\diamond$ onto functions of $v_\diamond$, using a single map θ shared across folds.

Entropy numbers and integrals. To control empirical process remainders, we use maximal inequalities based on uniform entropy integrals. For a uniformly bounded function class $\mathcal{F}$, let $N(\epsilon, \mathcal{F}, L_2(P))$ denote the ϵ-covering number [Van Der Vaart and Wellner, 1996] of $\mathcal{F}$ relative to the $L^2(P)$ metric, and define the uniform entropy integral of $\mathcal{F}$ by

$$\mathcal{J}(\delta, \mathcal{F}) := \int_0^\delta \sup_Q \sqrt{\log N(\epsilon, \mathcal{F}, L_2(Q))}\, d\epsilon\,,$$

where the supremum is taken over all discrete probability distributions Q. In contrast to Van Der Vaart and Wellner [1996], we do not define the uniform entropy integral relative to

an envelope function. We may do so because all function classes we consider are uniformly bounded; thus, any uniformly bounded envelope would change the uniform entropy integral in Van Der Vaart and Wellner [1996] only by a multiplicative constant.

Function classes used to control empirical process remainders. We define relevant function classes that will be used to control empirical process remainders in our proofs. Let $M/2 < \infty$ be an upper bound on the nuisance functions and estimators in C5. Let $\mathcal{F}_{\text{iso}} \subset \{f : [-M, M] \to \mathbb{R} \mid f \text{ is monotone nondecreasing}\}$ consist of all univariate isotonic functions with supremum norm at most $M < \infty$. Let $M' < \infty$ be an upper bound on the total variation norms of the calibration functions $\{\theta_{n,j,\mu}, \theta_{n,j,\alpha} : j \in [J]\}$ in C6. Let $\mathcal{F}_{\text{TV}} \subset \{f : [-M, M] \to \mathbb{R} \mid f \text{ has finite total variation}\}$ consist of all univariate functions with total variation norm at most $3M'$ and supremum norm at most $M < \infty$.

For $j \in [J]$, define the function classes

$$\mathcal{S}_{n,j} := \{f \circ \mu_{n,j} \,|\, f \in \mathcal{F}_{\text{TV}} \cup \mathcal{F}_{\text{iso}}\} \quad \text{and} \quad \mathcal{R}_{n,j} := \{g \circ \alpha_{n,j} \,|\, g \in \mathcal{F}_{\text{TV}} \cup \mathcal{F}_{\text{iso}}\}.$$

By C6 and C5, we will show that $r^*_{n,j} = \Pi_{\mu^*_{n,j}} \Delta_{\alpha,n,j}$ and $s^*_{n,j} = \Pi_{\alpha^*_{n,j}} \Delta_{\mu,n,j}$ can be expressed as one-dimensional transformations of $\mu_{n,j}$ and $\alpha_{n,j}$, respectively, whose total variation norms are bounded by $3M'$. Hence, with probability tending to one, $\mu^*_{n,j} = f_n \circ \mu_{n,j}$ and $s^*_{n,j}$ belong to $\mathcal{S}_{n,j}$. Similarly, $\alpha^*_{n,j} = g_n \circ \alpha_{n,j}$ and $r^*_{n,j}$ belong to $\mathcal{R}_{n,j}$. Note that $\mathcal{S}_{n,j}$ and $\mathcal{R}_{n,j}$ are fixed function classes conditional on the j-th training set $\mathcal{D}_n \backslash \mathcal{C}_j$.

Conventions for proofs. In our proofs, we may, without loss of generality, treat the bounds in C5 (and D1) as holding almost surely rather than only with probability tending to one. This is justified because our results are stated in terms of stochastic order (e.g., O_p and o_p), which is unaffected by restricting attention to events whose probabilities tend to one. Formally, we work on the event

$$\mathcal{B}_n := \Big\{\|s^*_{n,j}\|_\infty \le 2M,\ \|r^*_{n,j}\|_\infty \le 2M,\ \|\alpha^*_{n,j}\|_\infty \le 2M,\ \|\mu^*_{n,j}\|_\infty \le 2M,\ \forall j \in [J]\Big\},$$

which satisfies $P_0(\mathcal{B}_n) \to 1$ by C5. Since all subsequent bounds are derived on $\mathcal{B}_n$, we suppress the indicator $1\{\mathcal{B}_n\}$ in the notation and interpret stochastic order statements (e.g., $O_p(\cdot)$, $o_p(\cdot)$) as holding under P_0 on events whose probabilities tend to one. For two quantities x and y, we use the expression $x \lesssim y$ to mean that x is upper bounded by y times a universal constant that may only depend on global constants that appear in C5. We denote by $\mathcal{I}_Y$ the coordinate projection map $z = (w, a, y) \mapsto y$; for example, $P_0\{\mathcal{I}_Y - \mu_0\} = E_0[Y - \mu_0(A, W)]$.

A.3.2 Bias expansions for cross-fitted one-step estimators

We begin by deriving a generic error expansion for cross-fitted one-step estimators, which will serve as the starting point for our analysis. We then derive a cross-fitted analogue of Theorem 2.3, which we will use to establish DRAL for our isocalibrated one-step estimator.

Let $\alpha_{n,\diamond} = \{\alpha_{n,j} : j \in [J]\}$ and $\mu_{n,\diamond} = \{\mu_{n,j} : j \in [J]\}$ be arbitrary fold-specific estimators of α_0 and μ_0, respectively, which may be uncalibrated. Define the associated one-step debiased estimator by

$$\tau_n := \frac{1}{J}\sum_{j=1}^{J} \psi_{n,j}(\mu_{n,j}) \ + \ \frac{1}{n}\sum_{i=1}^{n} \alpha_{n,j(i)}(A_i, W_i)\{Y_i - \mu_{n,j(i)}(A_i, W_i)\}.$$

The following expansion of the estimation error $\tau_n - \tau_0$ is a standard starting point for the analysis of doubly robust estimators.

Lemma A.1 (Standard doubly robust bias expansion)**.** *It holds that*

$$\tau_n - \tau_0 = (\overline{P}_n - \overline{P}_0)\left\{\phi_{0,\mu_{n,\diamond}} + D_{\mu_{n,\diamond},\alpha_{n,\diamond}}\right\} + \langle \alpha_{n,\diamond} - \alpha_0, \mu_0 - \mu_{n,\diamond}\rangle_{\overline{P}_0} + Rem_{\mu_{n,\diamond}}(\psi_{n,\diamond}, \psi_0).$$

where $Rem_{\mu_{n,\diamond}}(\psi_{n,\diamond}, \psi_0) := \overline{\psi}_n(\mu_{n,\diamond}) - \overline{\psi}_0(\mu_{n,\diamond}) - \overline{P}_n\phi_{0,\mu_{n,\diamond}}$.

Proof. We can write

$$\frac{1}{J}\sum_{j=1}^{J}\{\psi_{n,j}(\mu_{n,j}) - \psi_0(\mu_{n,j})\} = (\overline{P}_n - \overline{P}_0)\phi_{0,\mu_{n,\diamond}} + Rem_{\mu_{n,\diamond}}(\psi_{n,\diamond}, \psi_0),$$

where

$$Rem_{\mu_{n,\diamond}}(\psi_{n,\diamond}, \psi_0) := \frac{1}{J}\sum_{j=1}^{J}\left\{\psi_{n,j}(\mu_{n,j}) - \psi_0(\mu_{n,j}) - (P_{n,j} - P_0)\phi_{0,\mu_{n,j}}\right\}.$$

The estimation error of τ_n can be decomposed as

$$\begin{aligned}\tau_n - \tau_0 &= \frac{1}{J}\sum_{j=1}^{J}\psi_{n,j}(\mu_{n,j}) + \overline{P}_n D_{\mu_{n,\diamond},\alpha_{n,\diamond}} - \psi_0(\mu_0)\\ &= (\overline{P}_n - \overline{P}_0)\left\{\phi_{0,\mu_{n,\diamond}} + D_{\mu_{n,\diamond},\alpha_{n,\diamond}}\right\} + \overline{P}_0 D_{\mu_{n,\diamond},\alpha_{n,\diamond}} + \frac{1}{J}\sum_{j=1}^{J}\{\psi_0(\mu_{n,j}) - \psi_0(\mu_0)\} + Rem_{\mu_{n,\diamond}}(\psi_{n,\diamond}, \psi_0).\end{aligned}$$

By the Riesz representation $\psi_0(\cdot) = \langle \alpha_0, \cdot\rangle_{P_0}$, we have $\psi_0(\mu_{n,j}) - \psi_0(\mu_0) = \langle \alpha_0, \mu_{n,j} - \mu_0\rangle_{P_0}$. Moreover, by the law of iterated expectations and $E_0[Y \mid A, W] = \mu_0(A, W)$,

$$\begin{aligned}\overline{P}_0 D_{\mu_{n,\diamond},\alpha_{n,\diamond}} &= \frac{1}{J}\sum_{j=1}^{J} E_0\left[\alpha_{n,j}(A,W)\{Y - \mu_{n,j}(A,W)\}\right]\\ &= \frac{1}{J}\sum_{j=1}^{J} E_0\left[\alpha_{n,j}(A,W)\{\mu_0(A,W) - \mu_{n,j}(A,W)\}\right]\\ &= \langle \alpha_{n,\diamond}, \mu_0 - \mu_{n,\diamond}\rangle_{\overline{P}_0}.\end{aligned}$$

Hence,

$$\overline{P}_0 D_{\mu_{n,\diamond},\alpha_{n,\diamond}} + \frac{1}{J}\sum_{j=1}^{J}\{\psi_0(\mu_{n,j}) - \psi_0(\mu_0)\} = \langle \alpha_{n,\diamond} - \alpha_0, \mu_0 - \mu_{n,\diamond}\rangle_{\overline{P}_0},$$

which yields the claimed expansion. □

The next lemma gives regression- and Riesz-based decompositions that characterize the leading-order behavior of the cross-product remainder. It is a cross-fitted analogue of Theorem 2.3, with explicit expressions for the first-order bias terms, which vanish under suitable orthogonality conditions on the cross-fitted nuisance estimators. The required orthogonality conditions are

$$\overline{P}_n\, s_{n,\diamond}\{\mathcal{I}_Y - \mu_{n,\diamond}\} = 0 \quad \text{and} \quad \overline{\psi}_n(r_{n,\diamond}) - \overline{P}_n\big(r_{n,\diamond}\alpha_{n,\diamond}\big) = 0,$$

where, for each $j \in [J]$, $r_{n,j} := \big(\overline{\Pi}_{\alpha_{n,\diamond}}(\mu_0 - \mu_{n,\diamond})\big)_j$ and $s_{n,j} := \big(\overline{\Pi}_{\mu_{n,\diamond}}(\alpha_0 - \alpha_{n,\diamond})\big)_j$. We also recall that $A_{n,j}\colon z \mapsto s_{n,j}(a,w)\{y - \mu_{n,j}(a,w)\}$ and $B_{n,j} := \psi_0(r_0) + \phi_{0,r_{n,j}} - r_{n,j}\alpha_{n,j}$. Finally, let $\Delta_{\alpha,n,j} := \alpha_{n,j} - \alpha_0$ and $\Delta_{\mu,n,j} := \mu_{n,j} - \mu_0$.

Lemma A.2 (Expansion of cross-product remainder)**.** *We have the regression-based decomposition:*

$$\langle \alpha_0 - \alpha_{n,\diamond},\ \mu_{n,\diamond} - \mu_0 \rangle_{\overline{P}_0} = -\overline{P}_n A_{n,\diamond} + (\overline{P}_n - \overline{P}_0) A_{n,\diamond} + \left\langle (\Pi_{(\mu_{n,\diamond},\mu_0)} - \Pi_{\mu_{n,\diamond}})\Delta_{\alpha,n,\diamond},\ \mu_0 - \Pi_{\mu_{n,\diamond}}\mu_0 \right\rangle_{\overline{P}_0}.$$

Furthermore, we have the Riesz-based decomposition:

$$\begin{aligned}\langle \alpha_0 - \alpha_{n,\diamond},\ \mu_{n,\diamond} - \mu_0 \rangle_{\overline{P}_0} = &- \left\{ \overline{\psi}_n(r_{n,\diamond}) - \overline{P}_n\big(r_{n,\diamond}\alpha_{n,\diamond}\big) \right\} + (\overline{P}_n - \overline{P}_0) B_{n,\diamond} \\ &+ \left\langle (\Pi_{(\alpha_{n,\diamond},\alpha_0)} - \Pi_{\alpha_{n,\diamond}})\Delta_{\mu,n,\diamond},\ \alpha_0 - \Pi_{\alpha_{n,\diamond}}\alpha_0 \right\rangle_{\overline{P}_0} + Rem_{\psi_{n,\diamond}}(r_{n,\diamond}),\end{aligned}$$

where $Rem_{\psi_{n,\diamond}}(r_{n,\diamond}) := \overline{\psi}_n(r_{n,\diamond}) - \overline{\psi}_0(r_{n,\diamond}) - \overline{P}_n\phi_{0,r_{n,\diamond}}$.

Proof. Our proof largely follows the proof of Theorem 2.3 after a change of notation. Specifically, let $j_{\text{rv}} \in [J]$ be uniformly distributed and independent of the data $\mathcal{D}_n$. Then $\widetilde{Z} := (Z, j_{\text{rv}}) \sim \overline{P}_0$, since $Z \sim P_0$ and $P(j_{\text{rv}} = j) = 1/J$ for $j \in [J]$. Our proof follows that of Theorem 2.3 applied with the extended data structure $\widetilde{Z}$ with P_0 replaced by $\overline{P}_0$ and with $\mu_n := \mu_{n,\diamond}$ and $\alpha_n := \alpha_{n,\diamond}$, where we view $(w, a, j) \mapsto \mu_{n,j}(a, w)$ as a function on $\mathcal{A} \times \mathcal{W} \times [J]$ (and similarly for $\alpha_{n,j}$). Under this identification, the P_0-projection operator Π_v is replaced by the $\overline{P}_0$-projection operator $\overline{\Pi}_{v_\diamond}$.

We begin with the proof of the Riesz-based decomposition, as the regression-based decomposition will follow as a special case. Since $\alpha_0 - \alpha_{n,\diamond} \in L^2(\overline{P}_0)$ is a measurable function of $(\alpha_{n,\diamond}, \alpha_0)$, applying the projection $\overline{\Pi}_{(\alpha_{n,\diamond},\alpha_0)}$ yields

$$\langle \alpha_0 - \alpha_{n,\diamond},\ \mu_{n,\diamond} - \mu_0 \rangle_{\overline{P}_0} = \left\langle \alpha_0 - \alpha_{n,\diamond},\ \overline{\Pi}_{(\alpha_{n,\diamond},\alpha_0)}\Delta_{\mu,n,\diamond} \right\rangle_{\overline{P}_0}.$$

We decompose the preceding display as

$$\begin{aligned}\langle \alpha_0 - \alpha_{n,\diamond},\ \mu_{n,\diamond} - \mu_0 \rangle_{\overline{P}_0} &= \left\langle \alpha_0 - \alpha_{n,\diamond},\ \overline{\Pi}_{\alpha_{n,\diamond}}\Delta_{\mu,n,\diamond} \right\rangle_{\overline{P}_0} \\ &\quad + \left\langle (\overline{\Pi}_{(\alpha_{n,\diamond},\alpha_0)} - \overline{\Pi}_{\alpha_{n,\diamond}})(\alpha_0 - \alpha_{n,\diamond}),\ \Delta_{\mu,n,\diamond} - \overline{\Pi}_{\alpha_{n,\diamond}}\Delta_{\mu,n,\diamond} \right\rangle_{\overline{P}_0} \\ &= \left\langle \alpha_{n,\diamond} - \alpha_0,\ r_{n,\diamond} \right\rangle_{\overline{P}_0} - \left\langle (\overline{\Pi}_{(\alpha_{n,\diamond},\alpha_0)} - \overline{\Pi}_{\alpha_{n,\diamond}})\Delta_{\alpha,n,\diamond},\ \Delta_{\mu,n,\diamond} - \overline{\Pi}_{\alpha_{n,\diamond}}\Delta_{\mu,n,\diamond} \right\rangle_{\overline{P}_0},\end{aligned} \tag{A.3}$$

where $r_{n,\diamond} = -\overline{\Pi}_{\alpha_{n,\diamond}}\Delta_{\mu,n,\diamond}$. By the Riesz representation property for α_0 in (2.1), the first term on the right-hand side of (A.3) can be written as

$$\begin{aligned}\langle \alpha_{n,\diamond} - \alpha_0,\ r_{n,\diamond} \rangle_{\overline{P}_0} &= -\psi_0(r_{n,\diamond}) + \overline{P}_0\{\alpha_{n,\diamond} r_{n,\diamond}\} \\ &= -\overline{\psi}_n(r_{n,\diamond}) + \overline{P}_0\{\alpha_{n,\diamond} r_{n,\diamond}\} - \psi_0(r_{n,\diamond}) + \overline{\psi}_n(r_{n,\diamond}),\end{aligned}$$

where, in the second equality, we add and subtract the cross-averaged functional estimate

$$\overline{\psi}_n(r_{n,\diamond}) := \frac{1}{J}\sum_{j=1}^{J} \psi_{n,j}(r_{n,j}).$$

Rearranging gives

$$\langle \alpha_{n,\diamond} - \alpha_0,\ r_{n,\diamond} \rangle_{\overline{P}_0} = -(\overline{P}_n - \overline{P}_0)\{\alpha_{n,\diamond} r_{n,\diamond}\} + \left\{ \overline{\psi}_n(r_{n,\diamond}) - \psi_0(r_{n,\diamond}) \right\} + \left\{ \overline{P}_n(\alpha_{n,\diamond} r_{n,\diamond}) - \overline{\psi}_n(r_{n,\diamond}) \right\}.$$

Next, add and subtract $\overline{P}_n\phi_{r_{n,\diamond}}$ and use $\overline{P}_0\phi_{r_{n,\diamond}} = 0$ to obtain

$$\langle \alpha_{n,\diamond} - \alpha_0,\ r_{n,\diamond} \rangle_{\overline{P}_0} = (\overline{P}_n - \overline{P}_0)\{-\alpha_{n,\diamond} r_{n,\diamond} + \phi_{r_{n,\diamond}}\}$$

$$+\left\{\overline{\psi}_n(r_{n,\diamond}) - \psi_0(r_{n,\diamond}) - (\overline{P}_n - \overline{P}_0)\phi_{r_{n,\diamond}}\right\}$$
$$+\left\{\overline{P}_n(\alpha_{n,\diamond} r_{n,\diamond}) - \overline{\psi}_n(r_{n,\diamond}) - \overline{P}_n\phi_{r_{n,\diamond}}\right\}.$$

Recalling

$$B_{n,\diamond} := \psi_0(r_{n,\diamond}) - \alpha_{n,\diamond} r_{n,\diamond} + \phi_{r_{n,\diamond}}, \qquad Rem_{\psi_n}(r_{n,\diamond}) := \overline{\psi}_n(r_{n,\diamond}) - \psi_0(r_{n,\diamond}) - (\overline{P}_n - \overline{P}_0)\phi_{r_{n,\diamond}},$$

we may rewrite the preceding display as

$$\langle \alpha_{n,\diamond} - \alpha_0,\, r_{n,\diamond}\rangle_{\overline{P}_0} = (\overline{P}_n - \overline{P}_0)B_{n,\diamond} + Rem_{\psi_n}(r_{n,\diamond}) + \left\{\overline{P}_n(\alpha_{n,\diamond} r_{n,\diamond}) - \overline{\psi}_n(r_{n,\diamond}) - \overline{P}_n\phi_{r_{n,\diamond}}\right\}.$$

Substituting this display into (A.3) yields

$$\begin{aligned}\langle \alpha_0 - \alpha_{n,\diamond},\, \mu_{n,\diamond} - \mu_0\rangle_{\overline{P}_0} &= (\overline{P}_n - \overline{P}_0)B_{n,\diamond} + Rem_{\psi_n}(r_{n,\diamond}) + \left\{\overline{P}_n(\alpha_{n,\diamond} r_{n,\diamond}) - \overline{\psi}_n(r_{n,\diamond}) - \overline{P}_n\phi_{r_{n,\diamond}}\right\} \\ &\quad - \left\langle (\overline{\Pi}_{(\alpha_{n,\diamond},\,\alpha_0)} - \overline{\Pi}_{\alpha_{n,\diamond}})\Delta_{\alpha,n,\diamond},\, \Delta_{\mu,n,\diamond} - \overline{\Pi}_{\alpha_{n,\diamond}}\Delta_{\mu,n,\diamond}\right\rangle_{\overline{P}_0} \\ &= (\overline{P}_n - \overline{P}_0)B_{n,\diamond} + Rem_{\psi_n}(r_{n,\diamond}) + \left\{\overline{P}_n(\alpha_{n,\diamond} r_{n,\diamond}) - \overline{\psi}_n(r_{n,\diamond}) - \overline{P}_n\phi_{r_{n,\diamond}}\right\} \\ &\quad + \left\langle \alpha_0 - \overline{\Pi}_{\alpha_{n,\diamond}}\alpha_0,\, \Delta_{\mu,n,\diamond} - \overline{\Pi}_{\alpha_{n,\diamond}}\Delta_{\mu,n,\diamond}\right\rangle_{\overline{P}_0},\end{aligned} \tag{A.4}$$

where we use that

$$(\overline{\Pi}_{(\alpha_{n,\diamond},\,\alpha_0)} - \overline{\Pi}_{\alpha_{n,\diamond}})\Delta_{\alpha,n,\diamond} = \overline{\Pi}_{\alpha_{n,\diamond}}\alpha_0 - \alpha_0 = -(\alpha_0 - \overline{\Pi}_{\alpha_{n,\diamond}}\alpha_0).$$

The Riesz-based decomposition now follows.

As in the proof of Theorem 2.3, the regression-based decomposition follows as a special case. For the regression-based decomposition, we apply the Riesz-based decomposition with

$$\psi_0(\alpha) := E_0\{\alpha(A,W)\mu_0(A,W)\} \quad \text{and} \quad \psi_{n,j}(\alpha) := E_{P_{n,j}}\{\alpha(A,W)Y\}.$$

In particular, μ_0 is the Riesz representer of the linear functional $\psi_0 : \alpha \mapsto E_0\{\alpha(A,W)\mu_0(A,W)\}$. For this choice of $\psi_{n,j}$, we have $Rem_{\psi_n}(\alpha) = 0$ with $\widetilde{\phi}_\alpha(z) := \alpha(A,W)Y - \psi_0(\alpha)$. Under these definitions, B_0 and $B_{n,\diamond}$ coincide with A_0 and $A_{n,\diamond}$, respectively.

□

A.3.3 Sufficient conditions for a negligible remainder: cross-fitted version of Lemma 2.4

Recall

$$\overline{\gamma}_{\mu,0}(\mu^*_{n,\diamond}, \alpha^*_{n,\diamond}) = \left\langle \overline{\Pi}_{(\mu^*_{n,\diamond},\,\mu_0)}\Delta_{\alpha,n,\diamond} - \overline{\Pi}_{\mu^*_{n,\diamond}}\Delta_{\alpha,n,\diamond}\,,\, \mu_0 - \overline{\Pi}_{\mu^*_{n,\diamond}}\mu_0\right\rangle_{\overline{P}_0},$$
$$\overline{\gamma}_{\alpha,0}(\mu^*_{n,\diamond}, \alpha^*_{n,\diamond}) = \left\langle \overline{\Pi}_{(\alpha^*_{n,\diamond},\,\alpha_0)}\Delta_{\mu,n,\diamond} - \overline{\Pi}_{\alpha^*_{n,\diamond}}\Delta_{\mu,n,\diamond}\,,\, \alpha_0 - \overline{\Pi}_{\alpha^*_{n,\diamond}}\alpha_0\right\rangle_{\overline{P}_0}.$$

We give sufficient conditions for the following requirement.

BB1) *Controlled partial orthogonality:*

i) $\overline{\gamma}_{\mu,0}(\mu^*_{n,\diamond}, \alpha^*_{n,\diamond}) = O_p(\|\mu^*_{n,\diamond} - \mu_0\|^2_{\overline{P}_0})$,
ii) $\overline{\gamma}_{\alpha,0}(\mu^*_{n,\diamond}, \alpha^*_{n,\diamond}) = O_p(\|\alpha^*_{n,\diamond} - \alpha_0\|^2_{\overline{P}_0})$.

The first sufficient condition is as follows.

BB1* *Projection error coupling:*

i) $\big\|(\overline{\Pi}_{(\mu_{n,\diamond},\mu_0)} - \overline{\Pi}_{\mu_{n,\diamond}})\Delta_{\alpha,n,\diamond}\big\|_{\overline{P}_0} = O_p\big(\|\mu^*_{n,\diamond} - \mu_0\|_{\overline{P}_0}\big)$;
ii) $\big\|(\overline{\Pi}_{(\alpha_{n,\diamond},\alpha_0)} - \overline{\Pi}_{\alpha_{n,\diamond}})\Delta_{\mu,n,\diamond}\big\|_{\overline{P}_0} = O_p\big(\|\alpha^*_{n,\diamond} - \alpha_0\|_{\overline{P}_0}\big)$.

The second sufficient condition is as follows. In the condition, let j_{rv} be uniformly distributed on $\{1, 2, \ldots, J\}$, independently of $\{Z_1, \ldots, Z_n\}$. Let $\overline{P}_0$ denote the joint distribution of (Z, j_{rv}), and write $\overline{E}_0 := E_{\overline{P}_0}$.

BB1** *Lipschitz continuity:* Denoting $\mathcal{D}_n := \{Z_1, \ldots, Z_n\}$, assume that, for all sufficiently large n, one of the following hold with P_0^n-probability one:

i) the map
$$(\widehat{m}, m) \mapsto \overline{E}_0\{\Delta_{\alpha,n,j_{\text{rv}}}(A, W) \,|\, \mu_{n,j_{\text{rv}}}(A, W) = \widehat{m},\ \mu_0(A, W) = m,\ \mathcal{D}_n\}$$
is Lipschitz continuous with a uniformly bounded Lipschitz constant;

ii) the map
$$(\widehat{a}, a) \mapsto \overline{E}_0\{\Delta_{\mu,n,j_{\text{rv}}}(A, W) \,|\, \alpha_{n,j_{\text{rv}}}(A, W) = \widehat{a},\ \alpha_0(A, W) = a,\ \mathcal{D}_n\}$$
is Lipschitz continuous with a uniformly bounded Lipschitz constant.

Lemma A.3 (Sufficient conditions for BB1)**.** *Conditions BB1i and BB1ii are implied by BB1*i and BB1*ii, respectively. Moreover, BB1*i and BB1*ii are implied by BB1**i and BB1**ii, respectively.*

Proof. We prove the result by reduction to the non–cross-fitted version Lemma 2.4 via an extended data-structure construction.

Let $j_{\text{rv}} \in [J]$ be uniformly distributed and independent of the data $\mathcal{D}_n$, and define the extended observation $\widetilde{Z} := (Z, j_{\text{rv}})$. Then $\widetilde{Z} \sim \overline{P}_0$. Moreover, we may view $\mu^*_{n,\diamond}, \alpha^*_{n,\diamond}$ as functions on $\mathcal{A} \times \mathcal{W} \times [J]$ via $(a, w, j) \mapsto \mu^*_{n,j}(a, w)$ and $(a, w, j) \mapsto \alpha^*_{n,j}(a, w)$, and similarly for $\Delta_{\alpha,n,\diamond}$ and $\Delta_{\mu,n,\diamond}$. Under this identification, the $\overline{P}_0$-projection $\overline{\Pi}_{v_\diamond}$ coincides with the usual $L^2(\overline{P}_0)$ projection operator Π_v applied to the corresponding function v on the extended space.

With this notation, $\gamma^*_{\mu,n,0}$ and $\gamma^*_{\alpha,n,0}$ are exactly the quantities $\gamma_{\mu,n,0}$ and $\gamma_{\alpha,n,0}$ from the non–cross-fitted setting in Section 2.3.2, but computed under $\overline{P}_0$ and with (μ_n, α_n) replaced by $(\mu^*_{n,\diamond}, \alpha^*_{n,\diamond})$. Likewise, Conditions BB1*i and BB1*ii (respectively, BB1**i and BB1**ii) are precisely the consistency conditions from Lemma 2.4, again after replacing P_0 by $\overline{P}_0$ and (μ_n, α_n) by $(\mu^*_{n,\diamond}, \alpha^*_{n,\diamond})$. Therefore, applying Lemma 2.4 on the extended space (with P_0 replaced by $\overline{P}_0$) yields that BB1*i implies BB1i, and BB1*ii implies BB1ii.

The second claim follows similarly from the second implication of Lemma 2.4. In particular, BB1**i implies that the projection
$$(\overline{\Pi}_{(\mu_{n,\diamond},\mu_0)}\Delta_{\alpha,n,\diamond})(A, W)$$
is a Lipschitz transformation of $(\mu_{n,\diamond}(A, W), \mu_0(A, W))$, and BB1**ii implies that the projection
$$(\overline{\Pi}_{(\alpha_{n,\diamond},\alpha_0)}\Delta_{\mu,n,\diamond})(A, W)$$
is a Lipschitz transformation of $(\alpha_{n,\diamond}(A, W), \alpha_0(A, W))$. Hence, BB1**i implies BB1*i, and BB1**ii implies BB1*ii. This completes the proof. □

A.3.4 Entropy bounds and local maximal inequalities for empirical processes

The following lemma controls the uniform entropy integral for the function classes $\mathcal{R}_{n,j}$ and $\mathcal{S}_{n,j}$, which contain $\mu^*_{n,j}$ and $\alpha^*_{n,j}$, respectively.

Lemma A.4. *For each $j \in [J]$, we have $\mathcal{J}(\delta, \mathcal{R}_{n,j}) + \mathcal{J}(\delta, \mathcal{S}_{n,j}) \lesssim \sqrt{\delta}$.*

Proof. We prove the claim for $\mathcal{S}_{n,j}$; the result for $\mathcal{R}_{n,j}$ follows by an identical argument. We have

$$\begin{aligned}
\mathcal{J}(\delta, \mathcal{S}_{n,j}) &= \int_0^\delta \sup_Q \sqrt{N(\varepsilon, \mathcal{S}_{n,j}, \|\cdot\|_Q)}\, d\varepsilon \\
&= \int_0^\delta \sup_Q \sqrt{N\left(\varepsilon, \mathcal{F}_{\mathrm{TV}} \cup \mathcal{F}_{\mathrm{iso}}, \|\cdot\|_{Q\circ\mu_{n,j}^{-1}}\right)}\, d\varepsilon \\
&\leq \int_0^\delta \sup_Q \sqrt{N\left(\varepsilon, \mathcal{F}_{\mathrm{TV}}, \|\cdot\|_{Q\circ\mu_{n,j}^{-1}}\right)}\, d\varepsilon + \int_0^\delta \sup_Q \sqrt{N\left(\varepsilon, \mathcal{F}_{\mathrm{iso}}, \|\cdot\|_{Q\circ\mu_{n,j}^{-1}}\right)}\, d\varepsilon \\
&\leq \mathcal{J}(\delta, \mathcal{F}_{\mathrm{TV}}) + \mathcal{J}(\delta, \mathcal{F}_{\mathrm{iso}}),
\end{aligned}$$

where $Q \circ \mu_{n,j}^{-1}$ denotes the push-forward probability measure of $\mu_{n,j}(W)$ conditional on the training sample $\mathcal{D}_n \setminus \mathcal{C}^{(j)}$. By standard metric entropy bounds for monotone functions, $\mathcal{J}(\delta, \mathcal{F}_{\mathrm{iso}}) \lesssim \sqrt{\delta}$ [Van Der Vaart and Wellner, 1996]. Moreover, since any function of finite variation can be written as the difference of two bounded monotone functions (e.g., Theorem 4 of Royden [1963]), it follows that

$$\mathcal{J}(\delta, \mathcal{F}_{\mathrm{TV}}) \lesssim \mathcal{J}(\delta, \mathcal{F}_{\mathrm{iso}}) \lesssim \sqrt{\delta}.$$

The result follows. □

The next lemma will be used to show that $r^*_{n,j}$ and $s^*_{n,j}$ are elements of $\mathcal{R}_{n,j} \cup \mathcal{S}_{n,j}$.

Lemma A.5. *Let $\widehat{\eta}, \eta_0 \in L^2(P_0)$ and let $\widehat{\eta}^* := \widehat{\theta} \circ \widehat{\eta}$ for a isotonic (monotone nondecreasing) piece-wise constant transformation $\widehat{\theta} : \mathbb{R} \to \mathbb{R}$ Suppose that there exists an $M < \infty$ such that the map $\widehat{\theta}_0 : \mathbb{R} \to \mathbb{R}$, defined such that $\widehat{\theta}_0(\widehat{\eta}(z)) := E_0[\eta_0(Z) \mid \widehat{\eta}(Z) = \widehat{\eta}(z), \mathcal{D}_n]$, has a total variation norm that is almost surely bounded by M. Then, the map $\widehat{\theta}^*_0$, defined such that $\widehat{\theta}^*_0(\widehat{\eta}(z)) = E_0[\eta_0(Z) \mid \widehat{\eta}^*(Z) = \widehat{\theta}(\widehat{\eta}(z)), \mathcal{D}_n]$, has its total variation, P_0-almost surely, bounded above by $3M$.*

Proof. This proof is a direct modification of the argument used to establish Lemma C.3 in Van Der Laan et al. [2023] (see also van der Laan et al. [2024b]), rewritten to match the present statement and notation.

Define the sets $B_t := \{z \in \mathbb{R} : \widehat{\theta}(z) = \widehat{\theta}(t)\}$. Since $\widehat{\theta}$ is monotone nondecreasing, each B_t is an interval of the form $B_t = \{z \in \mathbb{R} : a(t) \leq z < b(t)\}$ for some endpoints $a(t) \leq b(t)$ (possibly $a(t) = b(t)$). Let $\widehat{\theta}^*_0$ be as in the lemma statement. For any z,

$$\begin{aligned}
\widehat{\theta}^*_0(\widehat{\eta}(z)) &= E_0\left[\eta_0(Z) \mid \widehat{\eta}^*(Z) = \widehat{\theta}(\widehat{\eta}(z)), \mathcal{D}_n\right] \\
&= E_0\left[\eta_0(Z) \mid \widehat{\theta}(\widehat{\eta}(Z)) = \widehat{\theta}(\widehat{\eta}(z)), \mathcal{D}_n\right] \\
&= E_0\left[\eta_0(Z) \mid \widehat{\eta}(Z) \in B_{\widehat{\eta}(z)}, \mathcal{D}_n\right].
\end{aligned}$$

By the law of total expectation and the definition of $\widehat{\theta}_0$,

$$E_0\left[\eta_0(Z) \mid \widehat{\eta}(Z) \in B_{\widehat{\eta}(z)}, \mathcal{D}_n\right] = E_0\left[E_0\left[\eta_0(Z) \mid \widehat{\eta}(Z), \mathcal{D}_n\right] \,\middle|\, \widehat{\eta}(Z) \in B_{\widehat{\eta}(z)}, \mathcal{D}_n\right] = E_0\left[\widehat{\theta}_0(\widehat{\eta}(Z)) \,\middle|\, \widehat{\eta}(Z) \in B_{\widehat{\eta}(z)}, \mathcal{D}_n\right].$$

Thus, $\widehat{\theta}_0^*$ is obtained by "bin-averaging" the function $\widehat{\theta}_0$ over the partition induced by the level sets $\{B_t : t \in \mathbb{R}\}$ of the monotone step map $\widehat{\theta}$. Up to notation, this is exactly the setting of Lemma C.3 in Van Der Laan et al. [2023], which shows that such bin-averaging increases total variation by at most a factor of 3. For completeness, we prove this claim here.

Write the Jordan decomposition (Theorem 4, Section 5.2 of Royden, 1963) as $\widehat{\theta}_0 = \widehat{\theta}_0^+ - \widehat{\theta}_0^-$, where $\widehat{\theta}_0^+$ and $\widehat{\theta}_0^-$ are bounded, nondecreasing functions. Moreover, we may choose $\widehat{\theta}_0^+$ so that $\widehat{\theta}_0^+(\infty) - \widehat{\theta}_0^+(-\infty) = \|\widehat{\theta}_0\|_{TV}$. Since $\|\widehat{\theta}_0^-\|_{TV} = \|\widehat{\theta}_0 - \widehat{\theta}_0^+\|_{TV} \leq \|\widehat{\theta}_0\|_{TV} + \|\widehat{\theta}_0^+\|_{TV}$, it follows that $\|\widehat{\theta}_0^-\|_{TV} \leq 2\|\widehat{\theta}_0\|_{TV}$. Because $\widehat{\theta}$ is nondecreasing, by definition we have that if $\widehat{\theta}(t_1) < \widehat{\theta}(t_2)$, then $x_1 < x_2$ for any $x_1 \in B_{t_1}$ and $x_2 \in B_{t_2}$. It follows that both maps

$$t \mapsto E_0\left[\widehat{\theta}_0^+(\widehat{\eta}(Z)) \mid \widehat{\eta}(Z) \in B_t, \mathcal{D}_n\right], \qquad t \mapsto E_0\left[\widehat{\theta}_0^-(\widehat{\eta}(Z)) \mid \widehat{\eta}(Z) \in B_t, \mathcal{D}_n\right]$$

are nondecreasing (and bounded). Hence, by Theorem 4 of Royden [1963], their difference $\widehat{\theta}_0^* : t \mapsto E_0\left[\widehat{\theta}_0(\widehat{\eta}(Z)) \mid \widehat{\eta}(Z) \in B_t, \mathcal{D}_n\right]$ is of bounded variation. Moreover, the total variation of a monotone function equals the difference between its endpoint limits, and we have

$$\operatorname{ess\,inf}(\widehat{\theta}_0^+ \circ \widehat{\eta}) \leq E_0\left[\widehat{\theta}_0^+(\widehat{\eta}(Z)) \mid \widehat{\eta}(Z) \in B_t, \mathcal{D}_n\right] \leq \operatorname{ess\,sup}(\widehat{\theta}_0^+ \circ \widehat{\eta}),$$

and similarly for $\widehat{\theta}_0^- \circ \widehat{\eta}$. Consequently, the total variation norms of $t \mapsto E_0[\widehat{\theta}_0^+(\widehat{\eta}(Z)) \mid \widehat{\eta}(Z) \in B_t, \mathcal{D}_n]$ and $t \mapsto E_0[\widehat{\theta}_0^-(\widehat{\eta}(Z)) \mid \widehat{\eta}(Z) \in B_t, \mathcal{D}_n]$ are bounded by $\|\widehat{\theta}_0^+\|_{TV}$ and $\|\widehat{\theta}_0^-\|_{TV}$, respectively. Using sublinearity of the total variation norm, we conclude that the map $\widehat{\theta}_0^* : t \mapsto E_0\left[\widehat{\theta}_0(\widehat{\eta}(Z)) \mid \widehat{\eta}(Z) \in B_t, \mathcal{D}_n\right]$ has total variation norm bounded by

$$\|\widehat{\theta}_0^+\|_{TV} + \|\widehat{\theta}_0^-\|_{TV} \leq \|\widehat{\theta}_0\|_{TV} + 2\|\widehat{\theta}_0\|_{TV} = 3\|\widehat{\theta}_0\|_{TV}.$$

Since $\|\widehat{\theta}_0\|_{TV} \leq M$ almost surely by assumption, it follows that $\|\widehat{\theta}_0^*\|_{TV} \leq 3M$ P_0-almost surely.

□

The following lemma is a cross-averaged version of Lemma A.5, specialized to our isocalibrated cross-fitted nuisance estimators.

Lemma A.6 (Cross-averaged version of Lemma A.5)**.** *Under C6, there exists an $M < \infty$ such that, for all $n \in \mathbb{N}$, the functions $\theta_{n,\mu}^* : \mathbb{R} \to \mathbb{R}$ and $\theta_{n,\alpha}^* : \mathbb{R} \to \mathbb{R}$, defined by $\theta_{n,\mu}^* \circ \mu_{n,\diamond}^* = \overline{\Pi}_{\mu_{n,\diamond}^*} \Delta_{\alpha,n,\diamond}$ and $\theta_{n,\alpha}^* \circ \alpha_{n,\diamond}^* = \overline{\Pi}_{\alpha_{n,\diamond}^*} \Delta_{\mu,n,\diamond}$, have total variation norms that are almost surely bounded by $3M$.*

Proof. By symmetry, it suffices to prove the claim for $\theta_{n,\mu}^* \circ \mu_{n,\diamond}^* = \overline{\Pi}_{\mu_{n,\diamond}^*} \Delta_{\alpha,n,\diamond}$. For each $j \in [J]$, note that

$$\{\overline{\Pi}_{\mu_{n,\diamond}^*} \Delta_{\alpha,n,\diamond}\}_j(z) = E_{\overline{P}_0}\left[\Delta_{\alpha,n,j_{\mathrm{rv}}}(Z) \,\middle|\, \mu_{n,j_{\mathrm{rv}}}^*(Z) = \mu_{n,j}^*(z),\ \mathcal{D}_n\right],$$

where the expectation is taken over $\widetilde{Z} = (Z, j_{\mathrm{rv}}) \sim \overline{P}_0$, with j_{rv} uniformly distributed on $[J]$ and independent of Z.

Define $\widehat{\eta}(\widetilde{Z}) := \mu_{n,j_{\mathrm{rv}}}(Z)$, $\widehat{\eta}^*(\widetilde{Z}) := \mu^*_{n,j_{\mathrm{rv}}}(Z)$, and $\eta_0(\widetilde{Z}) := \Delta_{\alpha,n,j_{\mathrm{rv}}}(Z)$. By construction, $\widehat{\eta}^* = \widehat{\theta} \circ \widehat{\eta}$ for an isotonic, piecewise-constant map $\widehat{\theta}$.

Moreover, the map

$$\widehat{\theta}_0(t) := E_{\overline{P}_0}\Big[\eta_0(\widetilde{Z}) \;\Big|\; \widehat{\eta}(\widetilde{Z}) = t,\ \mathcal{D}_n\Big] = E_{\overline{P}_0}[\Delta_{\alpha,n,j_{\mathrm{rv}}}(Z) \mid \mu_{n,j_{\mathrm{rv}}}(Z) = t,\ \mathcal{D}_n]$$

has total variation norm almost surely bounded by M under C6. Therefore, applying Lemma A.5 with expectations taken under $\overline{P}_0$ yields that

$$\widehat{\theta}^*_0(t) := E_{\overline{P}_0}\Big[\eta_0(\widetilde{Z}) \;\Big|\; \widehat{\eta}^*(\widetilde{Z}) = t,\ \mathcal{D}_n\Big]$$

has total variation norm almost surely bounded by $3M$.

Setting $\theta^*_{n,\mu} := \widehat{\theta}^*_0$, we have

$$\{\overline{\Pi}_{\mu^*_{n,\diamond}} \Delta_{\alpha,n,\diamond}\}_j(z) = \theta^*_{n,\mu}\big(\mu^*_{n,j}(z)\big)\,,$$

and $\|\theta^*_{n,\mu}\|_{\mathrm{TV}} \le 3M$ almost surely, as claimed. □

The next lemma provides a local maximal inequality that we use to control the empirical process remainders that appear in our analysis.

Lemma A.7. *For each $j \in [J]$, let $\mathcal{H}_{n,j} = \mathcal{F}^4_{n,j}$ for a function class $\mathcal{F}_{n,j}$ that is deterministic conditional on the jth training set $\mathcal{D}_n \backslash \mathcal{C}^{(j)}$. Assume that $\mathcal{J}(\delta, \mathcal{F}_{n,j}) \le t_n(\delta)$ for some concave map $\delta \mapsto t_n(\delta)$ with $t_n(0) = 0$ and such that $\delta \mapsto t_n(\delta)/\delta$ is nonincreasing. For each $j \in [J]$, let $T_{n,j} : \mathcal{Z} \times \mathbb{R}^4 \to \mathbb{R}$ be deterministic conditional on $\mathcal{D}_n \setminus \mathcal{C}^{(j)}$ and uniformly bounded by some fixed $B < \infty$. Assume further that there exists $L < \infty$ such that, for each $z \in \mathcal{Z}$ and all $u, v \in \mathbb{R}^4$,*

$$\big|T_{n,j}(z,u) - T_{n,j}(z,v)\big| \le L\|u - v\|_2,$$

i.e., the map $T_{n,j}(z,\cdot)$ is L-Lipschitz uniformly in z. Then, for all $\delta > 0$,

$$E^n_0\left[\sup_{h_\diamond \in \mathcal{H}_{n,\diamond} : \|T_{n,\diamond}(\cdot, h_\diamond(\cdot))\|_{\overline{P}_0} \le \delta} (\overline{P}_n - \overline{P}_0)\, T_{n,\diamond}(\cdot, h_\diamond(\cdot))\right] \lesssim n^{-1/2} t_n(\delta)\left(1 + \frac{t_n(\delta)}{\sqrt{n}\,\delta^2}\right).$$

Furthermore, for any $\sqrt{n}\,\delta^2 > t_n(\delta)$,

$$E^n_0\left[\sup_{h_\diamond \in \mathcal{H}_{n,\diamond} : \|T_{n,\diamond}(\cdot, h_\diamond(\cdot))\|_{\overline{P}_0} \le \delta} \left(\overline{P}^{\#}_n - \overline{P}_n\right) T_{n,\diamond}(\cdot, h_\diamond(\cdot))\right] \lesssim n^{-1/2} t_n(\delta)\left(1 + \frac{t_n(\delta)}{\sqrt{n}\,\delta^2}\right),$$

where the implicit constant in "$\lesssim$" depends only on B, L, and the constants appearing in the stated conditions.

Proof. To establish the first bound, it suffices by the triangle inequality to show that, for each $j \in [J]$,

$$E^n_0\left[\sup_{h \in \mathcal{H}_{n,j} :\ \|T_{n,j}(\cdot, h(\cdot))\| \le \delta} (P_{n,j} - P_0)\, T_{n,j}(\cdot, h(\cdot))\right] \;\lesssim\; n^{-1/2}\, t_n(\delta)\left\{1 + \frac{t_n(\delta)}{\sqrt{n}\,\delta^2}\right\}.$$

The desired bound then follows since

$$E_0^n\left[\sup_{h_\diamond\in\mathcal{H}_{n,\diamond}:\ \|T_{n,\diamond}(\cdot,h_\diamond(\cdot))\|_{\overline{P}_0}\le\delta}\left(\overline{P}_n-\overline{P}_0\right)T_{n,\diamond}(\cdot,h_\diamond(\cdot))\right]\ \lesssim\ \max_{j\in[J]}E_0^n\left[\sup_{h\in\mathcal{H}_{n,j}:\ \|T_{n,j}(\cdot,h(\cdot))\|\le J\delta}(P_{n,j}-P_0)\,T_{n,j}(\cdot,h(\cdot))\right],$$

where we used that $\|T_{n,\diamond}(\cdot,h_\diamond(\cdot))\|_{\overline{P}_0}\le\delta$ implies $\|T_{n,j}(\cdot,h(\cdot))\|\le J\delta$ for each $j\in[J]$. Here, we also used that each fold has $\approx n/J$ folds with J fixed.

To this end, let $\mathcal{G}_{n,j}:=\{\,T_{n,j}(\cdot,h(\cdot)):h\in\mathcal{H}_{n,j}\,\}$. By Theorem 2.10.20 of Van Der Vaart and Wellner [1996] on preservation of entropy integrals under Lipschitz transformations and Lemma A.4, the class $\mathcal{G}_{n,j}$ is uniformly bounded and satisfies

$$\mathcal{J}(\delta,\mathcal{G}_{n,j})\lesssim\mathcal{J}(\delta,\mathcal{H}_{n,j})\lesssim 4\,\mathcal{J}(\delta,\mathcal{F}_{n,j})\lesssim t_n(\delta).$$

Hence, by Theorem 2.1 of van der Vaart and Wellner [2011], applied conditional on $\mathcal{D}_n\setminus\mathcal{C}^{(j)}$ and then averaged over $\mathcal{D}_n\setminus\mathcal{C}^{(j)}$ using the tower property,

$$E_0^n\left[\sup_{h\in\mathcal{H}_{n,j}:\ \|T_{n,j}(\cdot,h(\cdot))\|\le\delta}(P_{n,j}-P_0)\,T_{n,j}(\cdot,h(\cdot))\right]\ \lesssim\ n^{-1/2}\,t_n(\delta)\left(1+\frac{t_n(\delta)}{\sqrt{n}\,\delta^2}\right).$$

We obtain the second bound (the bootstrap case) by adapting the proof of Theorem 2.1 in van der Vaart and Wellner [2011]. Throughout, we work conditional on $\mathcal{D}_n\setminus\mathcal{C}^{(j)}$. Since the resulting bounds hold with absolute constants that do not depend on $\mathcal{D}_n\setminus\mathcal{C}^{(j)}$, averaging over $\mathcal{D}_n\setminus\mathcal{C}^{(j)}$ then yields the claimed result. To this end, note that $P_{n,j}^{\#}$ is the empirical distribution of $O(n/J)$ i.i.d. samples from $P_{n,j}$. Define the random radii

$$\delta_n:=\sup_{h\in\mathcal{H}_{n,j}:\ \|T_{n,j}(\cdot,h(\cdot))\|_{P_0}\le\delta}\|T_{n,j}(\cdot,h(\cdot))\|_{P_{n,j}}\quad\text{and}\quad\delta_n^{\#}:=\sup_{h\in\mathcal{H}_{n,j}:\ \|T_{n,j}(\cdot,h(\cdot))\|_{P_{n,j}}\le\delta_n}\|T_{n,j}(\cdot,h(\cdot))\|_{P_{n,j}^{\#}}.$$

By uniform boundedness of the class $\mathcal{G}_{n,j}$ and Equation (2.1) of van der Vaart and Wellner [2011] (a consequence of Dudley's entropy integral bound),

$$\begin{aligned}
&E_0^n\left[\sup_{h\in\mathcal{H}_{n,j}:\ \|T_{n,j}(\cdot,h(\cdot))\|_{P_0}\le J\delta}\left(P_{n,j}^{\#}-P_{n,j}\right)T_{n,j}(\cdot,h(\cdot))\,\middle|\,P_{n,j}\right]\\
&\lesssim n^{-1/2}\,E_0^n\left[\mathcal{J}(\delta_n^{\#},\mathcal{G}_{n,j})\,\middle|\,P_{n,j}\right]\\
&\lesssim n^{-1/2}\,E_0^n\left[t_n(\delta_n^{\#})\,\middle|\,P_{n,j}\right].
\end{aligned}$$

Since $\delta\mapsto t_n(\delta)$ is concave with $t_n(0)=0$ and $\delta\mapsto t_n(\delta)/\delta$ is nonincreasing, the proof of Theorem 2.1 of van der Vaart and Wellner [2011] implies that $\varepsilon\mapsto t_n(\sqrt{\varepsilon})$ is concave. Hence, by Jensen's inequality,

$$E_0^n\left[t_n(\delta_n^{\#})\,\middle|\,P_{n,j}\right]\ \le\ t_n\left(\widetilde{\delta}_n^{\#}\right),\qquad\widetilde{\delta}_n^{\#2}:=E_0^n\left[\delta_n^{\#2}\,\middle|\,P_{n,j}\right].$$

By the same argument and the tower property, averaging over $P_{n,j}$ yields

$$E_0^n\left[t_n(\delta_n^{\#})\right]\le E_0^n\left[t_n\left(\widetilde{\delta}_n^{\#}\right)\right]\le t_n\left(\widetilde{\delta}_n\right),\qquad\widetilde{\delta}_n^{2}:=E_0^n\left[\delta_n^{\#2}\right].$$

Thus,

$$E_0^n\left[\sup_{\substack{h\in\mathcal{H}_{n,j}:\\ \|T_{n,j}(\cdot,h(\cdot))\|_{P_0}\le J\delta}}\left(P_{n,j}^{\#}-P_{n,j}\right)T_{n,j}(\cdot,h(\cdot))\right]\lesssim n^{-1/2}t_n\left(\widetilde{\delta}_n\right).\tag{A.5}$$

We now further bound the right hand side. Equation (2.4) in the proof of Theorem 2.1 of van der Vaart and Wellner [2011] shows that

$$\widetilde{\delta}_n^{\#2} \lesssim \delta_n^2 + n^{-1/2}\, t_n(\widetilde{\delta}_n^{\#}).$$

See also Theorem 2.2 of van de Geer [2014] for a related result. Next, using again that $\varepsilon \mapsto t_n(\sqrt{\varepsilon})$ is concave and applying Jensen's inequality, we obtain

$$\begin{aligned}
\widetilde{\delta}_n^2 = E_0^n\left[\widetilde{\delta}_n^{\#2}\right] &\lesssim E_0^n\left[\delta_n^2\right] + n^{-1/2} E_0^n\left[t_n(\widetilde{\delta}_n^{\#})\right] \\
&\lesssim E_0^n\left[\delta_n^2\right] + n^{-1/2} t_n\left(\sqrt{E_0^n\left[\widetilde{\delta}_n^{\#2}\right]}\right) \\
&\lesssim E_0^n\left[\delta_n^2\right] + n^{-1/2} t_n\left(\widetilde{\delta}_n\right).
\end{aligned}$$

By Theorem 2.2 of van de Geer [2014], we have $E_0^n[\delta_n^2] \lesssim \delta^2$ for any $\delta > 0$ satisfying $\sqrt{n}\,\delta^2 > t_n(\delta)$. For such a δ, it follows that

$$\widetilde{\delta}_n^2 \lesssim \delta^2 + n^{-1/2} t_n\left(\widetilde{\delta}_n\right). \tag{A.6}$$

Finally, arguing as in the proof of Theorem 2.1 of van der Vaart and Wellner [2011], we apply Lemma 2.1 of that work to show that any $\widetilde{\delta}_n$ satisfying (A.6) also satisfies

$$t_n\left(\widetilde{\delta}_n\right) \leq t_n(\delta)\left(1 + \frac{t_n(\delta)}{\sqrt{n}\,\delta^2}\right).$$

Here we crucially use that $\delta \mapsto t_n(\delta)$ is concave with $t_n(0) = 0$, and that $\delta \mapsto t_n(\delta)/\delta$ is nonincreasing. Thus, applying this bound to (A.6), we conclude that

$$E_0^n\left[\sup_{\substack{h \in \mathcal{H}_{n,j}: \\ \|T_{n,j}(\cdot, h(\cdot))\|_{P_0} \leq \delta}} \left(P_{n,j}^{\#} - P_{n,j}\right) T_{n,j}(\cdot, h(\cdot)) \,\middle|\, P_{n,j}\right] \lesssim n^{-1/2} t_n(\delta)\left(1 + \frac{t_n(\delta)}{\sqrt{n}\,\delta^2}\right), \tag{A.7}$$

as desired.

□

Finally, the following lemma can be used to explicitly verify the locally uniform asymptotic linearity conditions for estimated functionals, such as Condition C3. We apply this lemma together with Condition D3 in the proof of Theorem 2.5.

Lemma A.8. *Suppose $\mathcal{A}$ is finite, and let $m : \mathcal{Z} \times \mathcal{H} \to \mathbb{R}$ be a functional that is linear in its second argument. Assume there exists $L \in (0, \infty)$ such that $|m(Z, h)| \leq L \max_{a \in \mathcal{A}} |h(a, W)|$ almost surely. Let $\nu \in \mathcal{H}$ be a fixed, uniformly bounded function, and let $\mathcal{F}$ be a class of functions from $\mathbb{R}$ to $\mathbb{R}$ with $\sup_{f \in \mathcal{F}} \|f\|_\infty \leq B$ and $\mathcal{J}(\infty, \mathcal{F}) < \infty$. Define the function classes $\mathcal{F}_\nu := \{f \circ \nu : f \in \mathcal{F}\}$ and $\mathcal{F}_{m,\nu} := \{m(\cdot, f) : f \in \mathcal{F}_\nu\}$. Then, there exists an $M < \infty$ depending only on L such that, for all $\delta > 0$, $\mathcal{J}(\delta, \mathcal{F}_\nu) \leq M\mathcal{J}(\delta, \mathcal{F})$ and $\mathcal{J}(\delta, \mathcal{F}_{m,\nu}) \leq M\mathcal{J}(\delta, \mathcal{F})$.*

Proof. Since $|m(Z, h)| \leq L \max_{a \in \mathcal{A}} |h(a, W)|$ almost surely, we have

$$N(\varepsilon, \mathcal{F}_{m,\nu}, L^2(Q)) \lesssim \sum_{a \in \mathcal{A}} N(\varepsilon, \mathcal{F}_\nu^{(a)}, L^2(Q)),$$

where we define the function class $\mathcal{F}_\nu^{(a)} := \{w \mapsto f(a,w) : f \in \mathcal{F}_\nu\}$ and use the finiteness of $\mathcal{A}$. Hence,

$$\mathcal{J}(\delta, \mathcal{F}_{m,\nu}) \lesssim \sum_{a\in\mathcal{A}} \mathcal{J}(\delta, \mathcal{F}_\nu^{(a)}).$$

By the definition of $\mathcal{F}_\nu$, we have $\mathcal{F}_\nu^{(a)} = \{w \mapsto f(\nu(a,w)) : f \in \mathcal{F}\}$. Now, for all $a \in \mathcal{A}$,

$$\begin{aligned}\mathcal{J}(\delta, \mathcal{F}_\nu^{(a)}) &= \int_0^\delta \sup_Q \sqrt{N(\varepsilon, \mathcal{F}_\nu^{(a)}, L^2(Q))}\, d\varepsilon \\ &= \int_0^\delta \sup_{Q_{a,\nu}} \sqrt{N(\varepsilon, \mathcal{F}, L^2(Q_{a,\nu}))}\, d\varepsilon \\ &= \mathcal{J}(\delta, \mathcal{F}),\end{aligned}$$

where $Q_{a,\nu}$ denotes the pushforward measure of the random variable $\nu(a, W)$. It follows that $\mathcal{J}(\delta, \mathcal{F}_{m,\nu}) \lesssim \mathcal{J}(\delta, \mathcal{F})$. By an identical argument, we also have $\mathcal{J}(\delta, \mathcal{F}_\nu) \lesssim \mathcal{J}(\delta, \mathcal{F})$. □

A.3.5 Dealing with empirical process remainders

In this section, we provide technical lemmas establishing the asymptotic negligibility of certain empirical process remainders. These lemmas rely on the following result, which gives sufficient conditions for these remainders to be asymptotically negligible.

Lemma A.9. *Assume C5 and C6. Let M be the uniform bound on Y and the nuisances in C5. Then, with P_0-probability tending to one, $s^*_{n,j}$ and $r^*_{n,j}$ are uniformly bounded by M for each $j = 1, \ldots, J$. For each $j \in [J]$, define*

$$\rho_{n,j}(z) := T_{n,j}\big(z, \mu^*_{n,j}, s^*_{n,j}, \alpha^*_{n,j}, r^*_{n,j}\big),$$

where $T_{n,j} : \mathcal{Z} \times [-2M, 2M]^4 \to \mathbb{R}$ satisfies the conditions of Lemma A.7. Suppose that $\|\rho_{n,\diamond}\|_{\overline{P}_0} = O_P(\varepsilon_n)$ for a deterministic sequence $\varepsilon_n \to 0$. Then

$$(\overline{P}_n - \overline{P}_0)\rho_{n,\diamond} = o_p(n^{-1/2}).$$

Moreover, if $\overline{P}_n$, $\mu^*_{n,j}$, $s^*_{n,j}$, $\alpha^*_{n,j}$, $r^*_{n,j}$, and $\rho_{n,j}$ are replaced by their bootstrap analogues $\overline{P}^\#_n$, $\mu^\#_{n,j}$, $s^\#_{n,j}$, $\alpha^\#_{n,j}$, $r^\#_{n,j}$, and $\rho^\#_{n,j}$, then

$$\left|(\overline{P}^\#_n - \overline{P}_n)\rho^\#_{n,\diamond}\right| + \left|(\overline{P}_n - \overline{P}_0)\rho^\#_{n,\diamond}\right| = o_p(n^{-1/2})$$

whenever $\|\rho^\#_{n,\diamond}\|_{\overline{P}_0} = O_P(\varepsilon_n)$.

Proof. First, note that $s^*_{n,j}$ is bounded by the uniform bound for $\alpha^*_{n,j} - \alpha_0$, since conditional expectations preserve essential suprema and, by C5, $\|\alpha^*_{n,j} - \alpha_0\|_\infty \le 2M$ with P_0–probability tending to one. Similarly, $\|r^*_{n,j}\|_\infty \le 2M$ with P_0–probability tending to one, since $r^*_{n,j}$ is bounded by the uniform bound for $\mu^*_{n,j} - \mu_0$. Therefore, with P_0–probability tending to one, we have

$$\big(z, \mu^*_{n,j}, s^*_{n,j}, \alpha^*_{n,j}, r^*_{n,j}\big) \in \mathcal{Z} \times [-2M, 2M]^4,$$

so $T_{n,j}\big(z, \mu^*_{n,j}, s^*_{n,j}, \alpha^*_{n,j}, r^*_{n,j}\big)$ is well-defined on this event. This establishes the first claim.

In what follows, we work on the event

$$\mathcal{B}_{n,j} := \Big\{\|s^*_{n,j}\|_\infty \le 2M,\ \|r^*_{n,j}\|_\infty \le 2M,\ \|\alpha^*_{n,j}\|_\infty \le 2M,\ \|\mu^*_{n,j}\|_\infty \le 2M\Big\},$$

which satisfies $P_0(\mathcal{B}_{n,j}) \to 1$ by C5. Since all subsequent bounds are derived on $\mathcal{B}_{n,j}$, we suppress the indicator $1\{\mathcal{B}_{n,j}\}$ in the notation and interpret stochastic order statements (e.g., $O_P(\cdot)$, $o_P(\cdot)$) as holding under P_0 on events whose probabilities tend to one.

We apply Lemma A.7 with $\mathcal{F}_{n,j} := \mathcal{S}_{n,j} \cup \mathcal{R}_{n,j}$, which contains $\mu^*_{n,j}$, $s^*_{n,j}$, $\alpha^*_{n,j}$, and $r^*_{n,j}$ with probability tending to one. This inclusion holds for $\mu^*_{n,j}$ and $\alpha^*_{n,j}$ because they are isotonic transformations of $\mu_{n,j}$ and $\alpha_{n,j}$, respectively. Similarly, by Lemma A.6, it holds for $r^*_{n,j}$ and $s^*_{n,j}$ because they are bounded-variation transformations of $\mu_{n,j}$ and $\alpha_{n,j}$, respectively. By Lemma A.4,

$$\mathcal{J}(\delta, \mathcal{F}_{n,j}) \lesssim \mathcal{J}(\delta, \mathcal{R}_{n,j}) + \mathcal{J}(\delta, \mathcal{S}_{n,j}) \lesssim t_n(\delta), \qquad \text{where } t_n(\delta) := \sqrt{\delta}.$$

Since $\|\rho_{n,\diamond}\|_{\overline{P}_0} = O_P(\varepsilon_n)$ with $\varepsilon_n \to 0$, there exists a deterministic sequence $\delta_n \to 0$ such that $\|\rho_{n,\diamond}\|_{\overline{P}_0} \lesssim \delta_n$ with probability tending to one. Let A_n denote the event $\{\|\rho_{n,\diamond}\|_{\overline{P}_0} \lesssim \delta_n\}$. On A_n,

$$\left|(\overline{P}_n - \overline{P}_0)\rho_{n,\diamond}\right| \le \sup_{\substack{h_\diamond \in \mathcal{H}_{n,\diamond}: \\ \|T_{n,\diamond}(\cdot, h_\diamond(\cdot))\|_{\overline{P}_0} \le \delta_n}} \left|(\overline{P}_n - \overline{P}_0)\, T_{n,\diamond}(\cdot, h_\diamond(\cdot))\right|.$$

Therefore, applying Lemma A.7 with $t_n(\delta) = \sqrt{\delta}$ yields

$$E_0^n\left[1(A_n) \sup_{\substack{h_\diamond \in \mathcal{H}_{n,\diamond}: \\ \|T_{n,\diamond}(\cdot, h_\diamond(\cdot))\|_{\overline{P}_0} \le \delta_n}} \left|(\overline{P}_n - \overline{P}_0)\, T_{n,\diamond}(\cdot, h_\diamond(\cdot))\right|\right] \lesssim n^{-1/2}\, t_n(\delta_n) \left\{1 + \frac{t_n(\delta_n)}{\sqrt{n}\,\delta_n^2}\right\}.$$

Since $t_n(\delta) = \sqrt{\delta}$, we have $t_n(\delta_n)\{1 + t_n(\delta_n)/(\sqrt{n}\,\delta_n^2)\} = \sqrt{\delta_n} + (n^{-1/2})\delta_n^{-3/2} = o(1)$ for any deterministic $\delta_n \downarrow 0$ satisfying $\sqrt{n}\,\delta_n^2 \to \infty$. Hence the displayed expectation is $o(n^{-1/2})$, and Markov's inequality implies

$$1(A_n)\,(\overline{P}_n - \overline{P}_0)\rho_{n,\diamond} = o_p(n^{-1/2}).$$

Because $P(A_n) \to 1$, it follows that $(\overline{P}_n - \overline{P}_0)\rho_{n,\diamond} = o_p(n^{-1/2})$.

The bootstrap claims follow by the same argument, using the second part of Lemma A.7 and noting that, conditional on $P_{n,j}$, $P^{\#}_{n,j}$ is the empirical distribution of $n_j \asymp n/J$ i.i.d. draws from $P_{n,j}$ (with J fixed). □

The following lemma establishes consistency of $s^*_{n,j}$ and $r^*_{n,j}$ in the appropriate regimes of C1 , as well as consistency and rate preservation of the bootstrapped nuisance estimators.

Lemma A.10 (Consistency of nuisance estimators and rate preservation under the bootstrap)**.** *Under C1 and C2, we have the following. Under C1ii ($\overline{\mu}_0 = \mu_0$), $\|s^*_{n,j} - s_0\| = o_p(1)$, and under C1i ($\overline{\alpha}_0 = \alpha_0$), $\|r^*_{n,j} - r_0\| = o_p(1)$.*

Furthermore, under Conditions C5, C3, and D3, whenever one part of C1 holds for $\alpha^*_{n,\diamond}$ and $\mu^*_{n,\diamond}$, the same part holds for $\alpha^{\#}_{n,\diamond}$ and $\mu^{\#}_{n,\diamond}$. In particular, under C1, we have $\|\alpha^{\#}_{n,\diamond} - \overline{\alpha}_0\|_{\overline{P}_0} = o_p(1)$ and $\|\mu^{\#}_{n,\diamond} - \overline{\mu}_0\|_{\overline{P}_0} = o_p(1)$. Moreover, under C1ii, $\|s^{\#}_{n,j} - s_0\| = o_p(1)$, and under C1i, $1\{\overline{\alpha}_0 = \alpha_0\}\,\|r^{\#}_{n,j} - r_0\| = o_p(1)$.

Proof. First, assume case C1ii in C1 ($\overline{\mu}_0 = \mu_0$). Then, the second part of C2ii yields

$$\|(\Pi_{\mu^*_{n,j}} - \Pi_{\mu_0})\Delta_{\alpha,n,j}\| = o_p(1).$$

Hence, by the triangle inequality,

$$\|s^*_{n,j} - s_0\| \le 1\{\overline{\mu}_0 = \mu_0\}\,\|(\Pi_{\mu^*_{n,j}} - \Pi_{\mu_0})\Delta_{\alpha,n,j}\| + 1\{\overline{\mu}_0 = \mu_0\}\,\|\Pi_{\mu_0}\{\alpha^*_{n,j} - \overline{\alpha}_0\}\| = o_p(1),$$

where we used the projection property

$$\|\Pi_{\mu_0}\{\alpha^*_{n,j} - \overline{\alpha}_0\}\| \le \|\alpha^*_{n,j} - \overline{\alpha}_0\| = o_p(1).$$

Next, assume case C1i ($\overline{\alpha}_0 = \alpha_0$). A symmetric argument applying C2i shows that

$$1\{\overline{\alpha}_0 = \alpha_0\}\,\|r^*_{n,j} - r_0\| = o_p(1).$$

Under Conditions C5, C3, and D3, the second part of Theorem A.15 implies that

$$\|\alpha^{\#}_{n,\diamond} - \overline{\alpha}_0\|_{\overline{P}_0} \le \|\alpha^*_{n,\diamond} - \overline{\alpha}_0\|_{\overline{P}_0} + O_p(n^{-1/3});$$
$$\|\mu^{\#}_{n,\diamond} - \overline{\mu}_0\|_{\overline{P}_0} \le \|\mu^*_{n,\diamond} - \overline{\mu}_0\|_{\overline{P}_0} + O_p(n^{-1/3}).$$

Hence, by C1,

$$\|\alpha^{\#}_{n,\diamond} - \overline{\alpha}_0\|_{\overline{P}_0} = o_p(1); \quad \|\mu^{\#}_{n,\diamond} - \overline{\mu}_0\|_{\overline{P}_0} = o_p(1).$$

Moreover, under C1i, we have $\|\alpha^{\#}_{n,\diamond} - \alpha_0\|_{\overline{P}_0} = o_p(n^{-1/4})$, whereas under C1ii, we have $\|\mu^{\#}_{n,\diamond} - \mu_0\|_{\overline{P}_0} = o_p(n^{-1/4})$. Thus, whenever one part of C1 holds for $\alpha^*_{n,\diamond}$ and $\mu^*_{n,\diamond}$, the same part holds for $\alpha^{\#}_{n,\diamond}$ and $\mu^{\#}_{n,\diamond}$.

Finally, applying the same argument as in the nonbootstrap case, we obtain $1\{\overline{\mu}_0 = \mu_0\}\,\|s^{\#}_{n,j} - s_0\| = o_p(1)$ and $1\{\overline{\alpha}_0 = \alpha_0\}\,\|r^{\#}_{n,j} - r_0\| = o_p(1)$ under their respective cases in C1, where we use the projection error bound in D2.

□

Lemma A.11 (Negligible empirical process remainders)**.** *Assume A3-C6 hold. Then, it holds that*

$$(\overline{P}_n - \overline{P}_0)\{D_{\mu^*_{n,\diamond},\alpha^*_{n,\diamond}} - D_{\overline{\mu}_0,\overline{\alpha}_0}\} = o_p(n^{-1/2}).$$

Moreover, under the setup of Theorem 2.5 and provided conditions D1-D3 also hold, it holds that

$$(\overline{P}^{\#}_n - \overline{P}_0)\{D_{\mu^{\#}_{n,\diamond},\alpha^{\#}_{n,\diamond}} - D_{\overline{\mu}_0,\overline{\alpha}_0}\} = o_p(n^{-1/2}).$$

Proof. Condition C1 implies that $\|\mu^*_{n,\diamond} - \overline{\mu}_0\|_{\overline{P}_0} = o_p(1)$ and $\|\alpha^*_{n,\diamond} - \overline{\alpha}_0\|_{\overline{P}_0} = o_p(1)$. Moreover, by the uniform boundedness in C5, the map $(\mu, \alpha) \mapsto D_{\mu,\alpha}$ is Lipschitz as a function into $L^2(\overline{P}_0)$. Hence,

$$\|D_{\mu^*_{n,\diamond},\alpha^*_{n,\diamond}} - D_{\overline{\mu}_0,\overline{\alpha}_0}\|_{\overline{P}_0} \lesssim \|\mu^*_{n,\diamond} - \overline{\mu}_0\|_{\overline{P}_0} + \|\alpha^*_{n,\diamond} - \overline{\alpha}_0\|_{\overline{P}_0} = o_p(1).$$

Furthermore, for each $j \in [J]$ and $z = (y, a, w) \in \mathcal{Z}$,

$$D_{\mu^*_{n,j},\alpha^*_{n,j}}(z) - D_{\overline{\mu}_0,\overline{\alpha}_0}(z) = T_{n,j}\big(z, \mu^*_{n,j}, 0, \alpha^*_{n,j}, 0\big),$$

where $T_{n,j}$ is the bounded Lipschitz transformation

$$T_{n,j}(z, \mu, s, \alpha, r) := \alpha(a, w)\{y - \mu(a, w)\}.$$

On the event in C5, we have $(z, \mu(a, w), s(a, w), \alpha(a, w), r(a, w)) \in \mathcal{Z} \times [-2M, 2M]^4$, so $T_{n,j}$ is uniformly bounded. It is also Lipschitz in (μ, α) on this bounded set because $T_{n,j}$ is formed from addition and multiplication, and multiplication is Lipschitz on bounded

domains. Therefore $T_{n,j}$ satisfies the conditions of Lemma A.9. Thus, by direct application of Lemma A.9, we have that

$$(\overline{P}_n - \overline{P}_0)\{D_{\mu^*_{n,\diamond},\alpha^*_{n,\diamond}} - D_{\overline{\mu}_0,\overline{\alpha}_0}\} = o_p(n^{-1/2}).$$

To prove the second part of this lemma, note that

$$\begin{aligned}(\overline{P}^{\#}_n - \overline{P}_0)\{D_{\mu^{\#}_{n,\diamond},\alpha^{\#}_{n,\diamond}} - D_{\overline{\mu}_0,\overline{\alpha}_0}\} &= (\overline{P}^{\#}_n - \overline{P}_n)\{D_{\mu^{\#}_{n,\diamond},\alpha^{\#}_{n,\diamond}} - D_{\overline{\mu}_0,\overline{\alpha}_0}\} \\ &\quad + (\overline{P}_n - \overline{P}_0)\{D_{\mu^{\#}_{n,\diamond},\alpha^{\#}_{n,\diamond}} - D_{\overline{\mu}_0,\overline{\alpha}_0}\}.\end{aligned}$$

By Lemma A.10, we have $\|\alpha^{\#}_{n,\diamond} - \overline{\alpha}_0\|_{\overline{P}_0} = o_p(1)$ and $\|\mu^{\#}_{n,\diamond} - \overline{\mu}_0\|_{\overline{P}_0} = o_p(1)$. Therefore, arguing as before and using D1, we obtain

$$\|D_{\mu^{\#}_{n,\diamond},\alpha^{\#}_{n,\diamond}} - D_{\overline{\mu}_0,\overline{\alpha}_0}\|_{\overline{P}_0} = o_p(1).$$

The second term on the right-hand side is $o_p(n^{-1/2})$ by the same argument as above, applying Lemma A.9 to the random function $z \mapsto D_{\mu^{\#}_{n,\diamond},\alpha^{\#}_{n,\diamond}}(z) - D_{\overline{\mu}_0,\overline{\alpha}_0}(z)$. The first term is also $o_p(n^{-1/2})$ by an analogous argument, using the bootstrap version of Lemma A.9.

□

For each $j \in [J]$, recall that $A^*_{n,j} : z \mapsto s^*_{n,j}(a,w)\{y - \mu^*_{n,j}(a,w)\}$ and $B^*_{n,j} := \psi_0(r_0) + \phi_{r^*_{n,j}} - r^*_{n,j}\alpha^*_{n,j}$. Also that $Rem_{\psi_{n,\diamond}}(r^*_{n,\diamond}) := \frac{1}{J}\sum_{j=1}^{J}\left\{\psi_{n,j}(r^*_{n,j}) - \psi_0(r^*_{n,j}) - P_{n,j}\phi_{0,r^*_{n,j}}\right\}$.

Lemma A.12 (Negligible cross-product empirical process remainders)**.** *Assume A1–C6. If C1ii ($\overline{\mu}_0 = \mu_0$), then*

$$(\overline{P}_n - P_0)\{A^*_{n,\diamond} - A_0\} = o_p(n^{-1/2}),$$

and if C1i ($\overline{\alpha}_0 = \alpha_0$), then

$$(\overline{P}_n - P_0)\{B^*_{n,\diamond} - B_0\} + Rem_{\psi_{n,\diamond}}(r^*_{n,\diamond}) = o_p(n^{-1/2}).$$

Moreover, under the setup of Theorem 2.5 and assuming D1–D3, the same conclusions hold with $\overline{P}_n$ replaced by $\overline{P}^{\#}_n$ and $A^*_{n,\diamond}, B^*_{n,\diamond}$ replaced by $A^{\#}_{n,\diamond}, B^{\#}_{n,\diamond}$.

Proof. The claim that

$$(\overline{P}_n - P_0)\{A^*_{n,\diamond} - A_0\} = o_p(n^{-1/2}) \quad \text{under C1ii}$$

follows from a modification of the proof of Lemma A.11. Specifically, we replace $\alpha^*_{n,j}$ with $s^*_{n,j}$ in the definition of $T_{n,j}$ and then apply Lemma A.9. This uses that

$$\|s^*_{n,j} - s_0\| = o_p(1) \quad \text{under C1ii},$$

which follows from Lemma A.10.

Next, we establish the claim that

$$(\overline{P}_n - P_0)\{B^*_{n,\diamond} - B_0\} = o_p(n^{-1/2})$$

under C1i. Note that

$$\frac{1}{J}\sum_{j=1}^{J}\left\{\psi_{n,j}(r^*_{n,j}) - \psi_0(r^*_{n,j}) - P_{n,j}\phi_{0,r_0}\right\} = o_p(n^{-1/2})$$

by C4. Hence,

$$
\begin{aligned}
(\overline{P}_n - P_0)\{B^*_{n,\diamond} - B_0\} + Rem_{\psi_{n,\diamond}}(r^*_{n,\diamond}) &= \frac{1}{J}\sum_{j=1}^{J}\Big\{(P_{n,j} - P_0)\{\phi_{0,r^*_{n,j}} - \phi_{0,r_0} + r_0\alpha_0 - r^*_{n,j}\alpha^*_{n,j}\}\Big\} \\
&\quad + \frac{1}{J}\sum_{j=1}^{J}\Big\{\psi_{n,j}(r^*_{n,j}) - \psi_0(r^*_{n,j}) - (P_{n,j} - P_0)\phi_{0,r^*_{n,j}}\Big\} \\
&= \frac{1}{J}\sum_{j=1}^{J}\{(P_{n,j} - P_0)\{r_0\alpha_0 - r^*_{n,j}\alpha^*_{n,j}\}\} \\
&\quad + \frac{1}{J}\sum_{j=1}^{J}\{\psi_{n,j}(r^*_{n,j}) - \psi_0(r^*_{n,j}) - (P_{n,j} - P_0)\phi_{0,r_0}\} \\
&= \frac{1}{J}\sum_{j=1}^{J}\{(P_{n,j} - P_0)\{r_0\alpha_0 - r^*_{n,j}\alpha^*_{n,j}\}\} + o_p(n^{-1/2}),
\end{aligned}
$$

where we used that $(P_{n,j} - P_0)\phi_{0,r_0} = P_{n,j}\phi_{0,r_0}$ since ϕ_{0,r_0} is mean zero. Arguing as in the proof of Lemma A.11, we may apply Lemma A.9 to conclude that

$$(\overline{P}_n - \overline{P}_0)\{r_0\alpha_0 - r^*_{n,j}\alpha^*_{n,j}\} = o_p(n^{-1/2}),$$

since, under C1i and C5, Lemma A.10 implies that

$$\|r_0\alpha_0 - r^*_{n,j}\alpha^*_{n,j}\|_{\overline{P}_0} \lesssim \|\alpha_0 - \alpha^*_{n,j}\|_{\overline{P}_0} + \|r_0 - r^*_{n,j}\|_{\overline{P}_0} = o_p(1).$$

The result now follows.

The bootstrap case follows along similar lines. Specifically, as in the proof of Lemma A.11 , we use the identity

$$(\overline{P}^{\#}_n - \overline{P}_0) = (\overline{P}_n - \overline{P}_0) + (\overline{P}^{\#}_n - \overline{P}_n),$$

where empirical process terms involving the first and second differences can be handled using the first and second parts of Lemma A.9, respectively. The bootstrap statements of Lemma A.10 show that

$$\|r^{\#}_{n,\diamond} - r_0\|_{\overline{P}_0} = o_p(1) \quad \text{under C1i},$$

and

$$\|s^{\#}_{n,\diamond} - s_0\|_{\overline{P}_0} = o_p(1) \quad \text{under C1ii}.$$

To show that

$$(\overline{P}^{\#}_n - P_0)\{B^{\#}_{n,\diamond} - B_0\} = o_p(n^{-1/2}),$$

we use that

$$\frac{1}{J}\sum_{j=1}^{J}\Big\{\psi_{n,j}(r^{\#}_{n,j}) - \psi_0(r^{\#}_{n,j}) - P_{n,j}\phi_{0,r_0}\Big\} + \frac{1}{J}\sum_{j=1}^{J}(P_{n,j} - P_0)\{m(\cdot, r^{\#}_{n,j}) - m(\cdot, r_0)\} = o_p(n^{-1/2})$$

by Lemma A.13 (proved next).

□

The following lemma can be used to explicitly verify the locally uniform asymptotic linearity condition in C4 for the estimated functional ψ_n when $\psi_n(\mu) = \frac{1}{n}\sum_{i=1}^{n} m(Z_i, \mu)$ and $\psi_0(\mu) = E_0\{m(Z, \mu)\}$, so that $\varphi_{0,\mu} = m(\cdot, \mu) - \psi_0(\mu)$ (the setup of Section 2.5.3).

Lemma A.13. *Consider the setup of Section 2.5.3, where $\psi_0(h) = E_0\{m(Z, h)\}$. Assume*

the conditions of Theorem 2.5. Then, for $(f_{n,j}, f_0) \in \{(\mu_{n,j}, \overline{\mu}_0), (\mu^{\#}_{n,j}, \overline{\mu}_0)\}$,

$$\frac{1}{J}\sum_{j=1}^{J}(P_{n,j}-P_0)\{m(\cdot,f_{n,j})-m(\cdot,f_0)\} = o_p(n^{-1/2}) \quad \textit{and} \quad \frac{1}{J}\sum_{j=1}^{J}(P^{\#}_{n,j}-P_{n,j})\{m(\cdot,f_{n,j})-m(\cdot,f_0)\} = o_p(n^{-1/2}).$$

The same conclusions hold for $(f_{n,j}, f_0) \in \{(r_{n,j}, r_0), (r^{\#}_{n,j}, r_0)\}$ *under C1i (so that* $\overline{\alpha}_0 = \alpha_0$*).*

Proof. Lemma A.9 is not directly applicable because, under Condition D3, the map $\mu \mapsto m(\cdot, \mu)$ is only pointwise Lipschitz in a sup-norm sense. Specifically, the condition assumes that $\mathcal{A}$ is finite and that there exists a constant $L \in (0, \infty)$ such that, for all $\mu \in \mathcal{H}$,

$$|m(Z, \mu)| \leq L \max_{a \in \mathcal{A}} |\mu(a, W)| \qquad P_0\text{-almost surely.}$$

In what follows, we adapt the proof of Lemma A.9.

By D3, we have $\|m(\cdot, \mu)\| \lesssim \|\mu\|$ for all $\mu \in \mathcal{H}$. Therefore, by C1, $\|m(\cdot, \mu^*_{n,j}) - m(\cdot, \overline{\mu}_0)\| \lesssim \|\mu^*_{n,j} - \overline{\mu}_0\| = o_p(1)$. Under C1i, by Lemma A.10, we have $\|r^*_{n,j} - r_0\| = o_p(1)$. Hence, under C1i, $\|m(\cdot, r^*_{n,j}) - m(\cdot, r_0)\| \lesssim \|r^*_{n,j} - r_0\| = o_p(1)$. We conclude that, under C1i,

$$\|m(\cdot, \mu^*_{n,j}) - m(\cdot, \overline{\mu}_0)\| = o_p(1) \quad \text{and} \quad \|m(\cdot, r^*_{n,j}) - m(\cdot, r_0)\| = o_p(1).$$

A symmetric argument using D2 also shows that

$$\|m(\cdot, \mu^{\#}_{n,j}) - m(\cdot, \overline{\mu}_0)\| = o_p(1) \quad \text{and} \quad \|m(\cdot, r^{\#}_{n,j}) - m(\cdot, r_0)\| = o_p(1).$$

Define the function class $\mathcal{F}_{m,n,j} := \{m(\cdot, f) - m(\cdot, g) : f \in \mathcal{R}_{n,j} \cup \mathcal{S}_{n,j}, g \in \{\overline{\mu}_0, r_0\}\}$. Note that $m(\cdot, \mu^*_{n,j}) - m(\cdot, \overline{\mu}_0)$, $m(\cdot, \mu^{\#}_{n,j}) - m(\cdot, \overline{\mu}_0)$, $m(\cdot, r^*_{n,j}) - m(\cdot, r_0)$, and $m(\cdot, r^{\#}_{n,j}) - m(\cdot, r_0)$ belong to $\mathcal{F}_{m,n,j}$ with probability tending to one. By Lemma A.8 and Lemma A.4, it holds that

$$\mathcal{J}(\delta, \mathcal{F}_{m,n,j}) \lesssim \mathcal{J}(\delta, \mathcal{R}_{n,j} \cup \mathcal{S}_{n,j}) \lesssim \sqrt{\delta}.$$

The proof of Lemma A.9 shows that $\mu^*_{n,j}$, $\alpha^*_{n,j}$, $r^*_{n,j}$, and $s^*_{n,j}$ lie in $\mathcal{R}_{n,j} \cup \mathcal{S}_{n,j}$ with probability tending to one. Consequently, the functions $m(\cdot, \mu^*_{n,j}) - m(\cdot, \overline{\mu}_0)$, $m(\cdot, \mu^{\#}_{n,j}) - m(\cdot, \overline{\mu}_0)$, $m(\cdot, r^*_{n,j}) - m(\cdot, r_0)$, and $m(\cdot, r^{\#}_{n,j}) - m(\cdot, r_0)$ belong to $\mathcal{F}_{m,n,j}$ with probability tending to one.

To complete the proof, we apply the local maximal inequalities in Lemma A.7 to each $\mathcal{F}_{m,n,j}$ and argue as in the proof of Lemma A.9 to obtain, for each $j \in [J]$,

$$(P_{n,j} - P_0)\{m(\cdot, \mu^*_{n,j}) - m(\cdot, \overline{\mu}_0)\} + (P_{n,j} - P_0)\{m(\cdot, r^*_{n,j}) - m(\cdot, r_0)\} = o_p(n^{-1/2}),$$

and

$$(P^{\#}_{n,j} - P_{n,j})\{m(\cdot, \mu^{\#}_{n,j}) - m(\cdot, \overline{\mu}_0)\} + (P^{\#}_{n,j} - P_{n,j})\{m(\cdot, r^{\#}_{n,j}) - m(\cdot, r_0)\} = o_p(n^{-1/2}).$$

The remaining bounds follow from symmetric arguments.

□

A.4 Convergence rates for isotonic-calibrated nuisance estimators

Let $\alpha^*_{n,\diamond} := \{\alpha^*_{n,j} : j \in [J]\}$ and $\mu^*_{n,\diamond} := \{\mu^*_{n,j} : j \in [J]\}$ denote the isotonic-calibrated nuisance estimators produced by Algorithm 1. Let $\overline{\alpha}_0$ and $\overline{\mu}_0$ be their respective (possibly misspecified) limits, as in Condition C1. In this section, we derive mean-squared error bounds for the calibrated estimators $\alpha^*_{n,j}$ and $\mu^*_{n,j}$ relative to α_0 and μ_0.

We work in the setting where the linear functional ψ_0 can be written as $\psi_0(\mu) = E_0\{m(Z,\mu)\}$ for a known map $m : \mathcal{Z} \times \mathcal{H} \to \mathbb{R}$, and we take ψ_n to be the empirical plug-in estimator $\psi_n(\mu) = n^{-1}\sum_{i=1}^n m(Z_i,\mu)$. The arguments below extend to more general estimators ψ_n, provided suitable regularity conditions hold for the associated second-order remainder.

We begin with a general result on convergence rates for isotonic-calibrated estimators of Riesz representers of bounded linear functionals. This framework includes both the outcome regression μ_0 and the Riesz representer α_0 as special cases. For $z \in \mathcal{Z}$ and $\nu \in \mathcal{H}$, define the Riesz loss [Chernozhukov et al., 2022c]

$$\ell(z,\nu) := \nu^2(a,w) - 2\,m(z,\nu),$$

where m is linear in its second argument. Let $\{\nu_{n,j} : j \in [J]\}$ be cross-fitted estimators of the population risk minimizer

$$\nu_0 := \arg\min_{\nu\in\mathcal{H}} P_0\,\ell(\cdot,\nu).$$

Write $\nu_{n,\diamond} := \{\nu_{n,j} : j \in [J]\}$, and define the calibrated cross-fitted estimators $\nu^*_{n,\diamond} := h_n \circ \nu_{n,\diamond}$ and $\nu^{\#}_{n,\diamond} := h^{\#}_n \circ \nu_{n,\diamond}$, where the isotonic regression maps h_n, $h^{\#}_n$, and $h_{n,0}$ are given by

$$h_n := \arg\min_{h\in\mathcal{F}_{\text{iso}}} \sum_{i=1}^n \ell\big(Z_i,\ h\circ\nu_{n,j(i)}\big)\,,$$

$$h^{\#}_n := \arg\min_{h\in\mathcal{F}_{\text{iso}}} \sum_{i=1}^n \ell\Big(Z^{\#}_i,\ h\circ\nu_{n,j(i)}\Big)\,,$$

$$h_{n,0} := \arg\min_{h\in\mathcal{F}_{\text{iso}}} \sum_{j=1}^J E_0[\ell(Z,\ h\circ\nu_{n,j})] \;=\; \arg\min_{h\in\mathcal{F}_{\text{iso}}} \big\|h\circ\nu_{n,\diamond} - \nu_0\big\|_{\overline{P}_0}.$$

We impose the following conditions.

(F1) With probability tending to one, h_n, $h^{\#}_n$, and $h_{n,0}$ lie in a fixed, uniformly bounded subset of $\mathcal{F}_{\text{iso}}$.

(F2) There exists $L < \infty$ such that $|E_0\{m(Z,h)\}| \le L\|h\|_{P_0}$ for all $h \in \mathcal{H}$.

(F3) The action space $\mathcal{A}$ is finite, and there exists $L \in (0,\infty)$ such that for all $\mu \in \mathcal{H}$, $|m(Z,\mu)| \le L\max_{a\in\mathcal{A}}|\mu(a,W)|$ almost surely.

Theorem A.14. *Assume F1-F3. Then,* $\|h_n\circ\nu_{n,\diamond} - h_{n,0}\circ\nu_{n,\diamond}\|_{\overline{P}_0} = O_p(n^{-1/3})$. *Hence,*

$$\|h_n\circ\nu_{n,\diamond} - \nu_0\|_{\overline{P}_0} \le \inf_{h\in\mathcal{F}_{iso}} \|h\circ\nu_{n,\diamond} - \nu_0\| + O_p(n^{-1/3})$$

Moreover, the bootstrap calibrated nuisance estimators satisfy, for any function $\overline{\nu}_0 \in L^2(P_0)$ *and as* $n\to\infty$, *the following root mean square error bounds:*

$$\|h^{\#}_n\circ\nu_{n,\diamond} - \overline{\nu}_0\|_{\overline{P}_0} \le \|h_n\circ\nu_{n,\diamond} - \overline{\nu}_0\|_{\overline{P}_0} + O_p(n^{-1/3}).$$

Proof. Let $\widetilde{\mathcal{F}}_{iso}$ be a uniformly bounded subset of $\mathcal{F}_{iso}$ that contains h_n, $h^{\#}_n$, and $h_{n,0}$ with probability tending to one (which exists by F1). Define $\mathcal{F}_{n,j} := \{f\circ\nu_{n,j} : f \in \widetilde{\mathcal{F}}_{iso}\}$. Since

h_n and $h_{n,0}$ fall in $\mathcal{F}_{iso}$ with probability tending to one by F1, it holds that $\nu^*_{n,j} = h_n \circ \nu_{n,j}$ and $h_{n,0} \circ \nu_{n,j}$ fall in $\mathcal{F}_{n,j}$ with probability tending to one.

We first show that the population risk induced by the Riesz loss is quadratic and strongly convex. By Assumption F2, the map $\nu \mapsto P_0 m(\cdot, \nu)$ is a bounded linear functional on $\mathcal{H}$. Moreover, ℓ is the corresponding Riesz loss [Chernozhukov et al., 2022c], so the population minimizer $\nu_0 := \arg\min_{\nu \in \mathcal{H}} P_0 \ell(\cdot, \nu)$ is its (nonparametric) Riesz representer. In particular, the Riesz representation property yields $P_0 m(\cdot, \nu) = P_0(\nu\nu_0)$ for all $\nu \in \mathcal{H}$, and hence

$$P_0 \ell(\cdot, \nu) = P_0\{\nu^2 - 2m(\cdot, \nu)\} = P_0\{\nu^2 - 2\nu\nu_0\}.$$

Now fix a uniformly bounded, closed, convex class $\mathcal{F} \subseteq \mathcal{H}$, and let $\nu_{0,\mathcal{F}} := \arg\min_{\nu \in \mathcal{F}} P_0 \ell(\cdot, \nu)$. Since $\nu \mapsto P_0 \ell(\cdot, \nu)$ is differentiable and convex, the first-order optimality condition at $\nu_{0,\mathcal{F}}$ gives, for all $\nu \in \mathcal{F}$,

$$\frac{d}{dt} P_0 \ell\big(\cdot, \nu_{0,\mathcal{F}} + t(\nu - \nu_{0,\mathcal{F}})\big)\Big|_{t=0} = 2\, P_0\{(\nu_{0,\mathcal{F}} - \nu_0)(\nu - \nu_{0,\mathcal{F}})\} \geq 0,$$

equivalently $P_0\{(\nu_{0,\mathcal{F}} - \nu_0)(\nu - \nu_{0,\mathcal{F}})\} \geq 0$. Using the quadratic form above, we then expand the excess risk:

$$\begin{aligned} P_0\{\ell(\cdot, \nu) - \ell(\cdot, \nu_{0,\mathcal{F}})\} &= P_0\{(\nu - \nu_{0,\mathcal{F}})(\nu + \nu_{0,\mathcal{F}} - 2\nu_0)\} \\ &= \|\nu - \nu_{0,\mathcal{F}}\|^2_{P_0} + 2\, P_0\{(\nu_{0,\mathcal{F}} - \nu_0)(\nu - \nu_{0,\mathcal{F}})\} \\ &\geq \|\nu - \nu_{0,\mathcal{F}}\|^2_{P_0}, \end{aligned}$$

which establishes the strong-convexity inequality

$$\|\nu - \nu_{0,\mathcal{F}}\|^2_{P_0} \leq P_0\{\ell(\cdot, \nu) - \ell(\cdot, \nu_{0,\mathcal{F}})\}.$$

We prove the convergence rates for the bootstrap estimators, since the rates for the non-bootstrap estimators follow by similar arguments. To this end, for $j \in [J]$, define the random rate $\delta^{\#}_{n,j} := \|h^{\#}_n \circ \nu_{n,j} - h_{n,0} \circ \nu_{n,j}\|_{P_0}$ and let $(\delta^{\#}_n)^2 := \frac{1}{J}\sum_{j=1}^{J} (\delta^{\#}_{n,j})^2 = \|h^{\#}_n \circ \nu_{n,\diamond} - h_{n,0} \circ \nu_{n,\diamond}\|^2_{\overline{P}_0}$. By convexity of $\mathcal{F}_{iso}$, the quadratic excess risk bound of the previous paragraph, and the fact that $h_{n,0}$ is a population risk minimizer, we have

$$(\delta^{\#}_n)^2 \lesssim \overline{P}_0 \ell(\cdot, h^{\#}_n \circ \nu_{n,\diamond}) - \overline{P}_0 \ell(\cdot, h_{n,0} \circ \nu_{n,\diamond}).$$

Furthermore, since $h^{\#}_n \circ \nu_{n,\diamond}$ is the bootstrap empirical risk minimizer, it holds that

$$\overline{P}^{\#}_n \Big\{ \ell(\cdot, h^{\#}_n \circ \nu_{n,\diamond}) - \ell(\cdot, h_{n,0} \circ \nu_{n,\diamond}) \Big\} \leq 0.$$

Combining the above two displays, we obtain

$$\begin{aligned} \overline{P}_0 \ell(\cdot, h^{\#}_n \circ \nu_{n,\diamond}) - \overline{P}_0 \ell(\cdot, h_{n,0} \circ \nu_{n,\diamond}) &= (\overline{P}_0 - \overline{P}^{\#}_n)\ell(\cdot, h^{\#}_n \circ \nu_{n,\diamond}) \\ &\quad + \overline{P}^{\#}_n \Big\{ \ell(\cdot, h^{\#}_n \circ \nu_{n,\diamond}) - \ell(\cdot, h_{n,0} \circ \nu_{n,\diamond}) \Big\} \\ &\quad - (\overline{P}_0 - \overline{P}^{\#}_n)\ell(\cdot, h_{n,0} \circ \nu_{n,\diamond}) \\ &\leq (\overline{P}_0 - \overline{P}^{\#}_n) \Big\{ \ell(\cdot, h^{\#}_n \circ \nu_{n,\diamond}) - \ell(\cdot, h_{n,0} \circ \nu_{n,\diamond}) \Big\}, \end{aligned}$$

and therefore

$$\begin{aligned} (\delta^{\#}_n)^2 &\leq (\overline{P}_0 - \overline{P}^{\#}_n) \Big\{ \ell(\cdot, h^{\#}_n \circ \nu_{n,\diamond}) - \ell(\cdot, h_{n,0} \circ \nu_{n,\diamond}) \Big\} \\ &\leq (\overline{P}_0 - \overline{P}_n) \Big\{ \ell(\cdot, h^{\#}_n \circ \nu_{n,\diamond}) - \ell(\cdot, h_{n,0} \circ \nu_{n,\diamond}) \Big\} \end{aligned}$$

$$
\begin{aligned}
&\quad + (\overline{P}_n - \overline{P}_n^{\#}) \left\{ \ell(\cdot, h_n^{\#} \circ \nu_{n,\diamond}) - \ell(\cdot, h_{n,0} \circ \nu_{n,\diamond}) \right\} \\
&\leq \frac{1}{J} \sum_{j=1}^{J} \left[\gamma_{n,j}(\delta_{n,j}^{\#}) + \gamma_{n,j}^{\#}(\delta_{n,j}^{\#}) \right],
\end{aligned}
$$

where we apply the triangle inequality and define

$$
\begin{aligned}
\gamma_{n,j}(\delta) &:= \sup_{\nu_1, \nu_2 \in \mathcal{F}_{n,j}:\ \|\nu_1 - \nu_2\| \leq \delta} \left| (P_0 - P_{n,j}) \left\{ \ell(\cdot, \nu_1) - \ell(\cdot, \nu_2) \right\} \right|, \\
\gamma_{n,j}^{\#}(\delta) &:= \sup_{\nu_1, \nu_2 \in \mathcal{F}_{n,j}:\ \|\nu_1 - \nu_2\| \leq \delta} \left| (P_{n,j} - P_{n,j}^{\#}) \left\{ \ell(\cdot, \nu_1) - \ell(\cdot, \nu_2) \right\} \right|.
\end{aligned}
$$

Moreover, since $\max_{j \in [J]} \delta_{n,j}^{\#} \leq J \delta_n^{\#}$, we have

$$
(\delta_n^{\#})^2 \leq \frac{1}{J} \sum_{j=1}^{J} \left[\gamma_{n,j}(J\delta_n^{\#}) + \gamma_{n,j}^{\#}(J\delta_n^{\#}) \right]. \tag{A.8}
$$

By Assumption F3 and the boundedness of $\mathcal{F}_{n,j}$, the function class $\{\ell(\cdot, \nu_1) - \ell(\cdot, \nu_2) : \nu_1, \nu_2 \in \mathcal{F}_{n,j}\}$ has a controlled uniform entropy integral, conditional on the training data for the nuisance estimators (so that the class is fixed given this training data). In particular, by preservation properties of the uniform entropy integral [Van Der Vaart and Wellner, 1996], the product class $\mathcal{F}_{n,j} \times \mathcal{F}_{n,j}$ has a uniform entropy integral bounded, up to a constant, by $\mathcal{J}(\delta, \mathcal{F}_{\nu_{n,j},\text{iso}}) + \mathcal{J}(\delta, \mathcal{F}_{\nu_{n,j},\text{iso}})$ for each $j \in [J]$. This quantity is itself, up to a constant, bounded above by $\mathcal{J}(\delta, \mathcal{F}_{\text{iso}})$ (see, e.g., the proof of Lemma A.4). Therefore, applying Lemma A.8 and following an argument identical to the proof of Lemma A.7 with $t_n(\delta) := \sqrt{\delta}$, we obtain, for any fixed $\delta > 0$, conditional on the training sample $\mathcal{T}_{n,j} := \mathcal{D}_n \setminus \mathcal{C}^{(j)}$,

$$
\begin{aligned}
E_0^n[\gamma_{n,j}(\delta) \mid \mathcal{T}_{n,j}] &\lesssim n^{-\frac{1}{2}} \sqrt{\delta} \left(1 + \frac{\sqrt{\delta}}{\sqrt{n}\,\delta^2} \right), \\
E_0^n\left[\gamma_{n,j}^{\#}(\delta) \mid \mathcal{T}_{n,j}\right] &\lesssim n^{-\frac{1}{2}} \sqrt{\delta} \left(1 + \frac{\sqrt{\delta}}{\sqrt{n}\,\delta^2} \right),
\end{aligned} \tag{A.9}
$$

where the bound on the right-hand side is deterministic with "$\lesssim$" holding for a fixed constant.

We are now ready to obtain a convergence rate bound for $\delta_n^{\#} = \|h_n^{\#} \circ \nu_{n,\diamond} - h_{n,0} \circ \nu_{n,\diamond}\|_{\overline{P}_0}$. To do so, we use a peeling argument similar to the proof of Theorem 3.2.5 in Van Der Vaart and Wellner [1996]. Proceeding with the proof and letting $\varepsilon_n \in \mathbb{R}$ be a deterministic rate to be set later and using that (A.8) holds with probability one

$$
\begin{aligned}
P\left(2^{k+1}\varepsilon_n \geq \delta_n^{\#} \geq 2^k \varepsilon_n\right) &= P\left(2^{k+1}\varepsilon_n \geq \delta_n^{\#} \geq 2^k \varepsilon_n,\ (\delta_n^{\#})^2 \leq \frac{1}{J} \sum_{j=1}^{J} \left[\gamma_{n,j}(J\delta_n^{\#}) + \gamma_{n,j}^{\#}(J\delta_n^{\#}) \right] \right) \\
&\leq P\left(2^{k+1}\varepsilon_n \geq \delta_n^{\#} \geq 2^k \varepsilon_n,\ 2^{2k}\varepsilon_n^2 \leq \frac{1}{J} \sum_{j=1}^{J} \left[\gamma_{n,j}(J2^{k+1}\varepsilon_n) + \gamma_{n,j}^{\#}(J2^{k+1}\varepsilon_n) \right] \right) \\
&\leq P\left(2^{2k}\varepsilon_n^2 \leq \frac{1}{J} \sum_{j=1}^{J} \left[\gamma_{n,j}(J2^{k+1}\varepsilon_n) + \gamma_{n,j}^{\#}(J2^{k+1}\varepsilon_n) \right] \right) \\
&\leq \sum_{j=1}^{J} E_0^n \left[P\left(2^{2k}\varepsilon_n^2 \leq \gamma_{n,j}(J2^{k+1}\varepsilon_n) + \gamma_{n,j}^{\#}(J2^{k+1}\varepsilon_n) \,\middle|\, \mathcal{T}_{n,j}\right) \right]
\end{aligned}
$$

$$\leq \sum_{j=1}^{J} E_0^n \left[\frac{E_0^n \left[\gamma_{n,j}(J2^{k+1}\varepsilon_n) \mid \mathcal{T}_{n,j} \right] + E_0^n \left[\gamma_{n,j}^{\#}(J2^{k+1}\varepsilon_n) \mid \mathcal{T}_{n,j} \right]}{2^{2k}\varepsilon_n^2} \right].$$

Taking $\varepsilon_n := n^{-1/3}$ above and applying (A.9), we can show, for some constant $C' > 0$, that

$$P\left(2^{k+1}n^{-1/3} \geq \delta_n^{\#} \geq 2^k n^{-1/3}\right) \lesssim \frac{n^{-1/2}\sqrt{2^{k+1}\varepsilon_n}\left(1 + \frac{\sqrt{2^{k+1}\varepsilon_n}}{\sqrt{n}\,\varepsilon_n^2}\right)}{2^{2k}\varepsilon_n^2} \leq C'\, 2^{-k}.$$

Therefore, for any $K > 0$,

$$P\left(\delta_n^{\#} \geq 2^K n^{-1/3}\right) = \sum_{k=K}^{\infty} P\left(2^{k+1}n^{-1/3} \geq \delta_n^{\#} \geq 2^k n^{-1/3}\right) \leq C' \sum_{k=K}^{\infty} 2^{-k},$$

and $\sum_{k=K}^{\infty} 2^{-k} \to 0$ as $K \to \infty$. Hence, for every $\varepsilon > 0$, there exists $K > 0$ such that $P(\delta_n^{\#} \geq 2^K n^{-1/3}) \leq \varepsilon$. By definition, $\delta_n^{\#} = O_p(n^{-1/3})$, and therefore

$$\|h_n^{\#} \circ \nu_{n,\diamond} - h_{n,0} \circ \nu_{n,\diamond}\|_{\overline{P}_0}^2 = O_p(n^{-2/3}).$$

An analogous argument for the non-bootstrap case yields

$$\|h_n \circ \nu_{n,\diamond} - h_{n,0} \circ \nu_{n,\diamond}\|_{\overline{P}_0}^2 = O_p(n^{-2/3}).$$

By the triangle inequality,

$$\|h_n^{\#} \circ \nu_{n,\diamond} - h_n \circ \nu_{n,\diamond}\|_{\overline{P}_0} \leq \|h_n^{\#} \circ \nu_{n,\diamond} - h_{n,0} \circ \nu_{n,\diamond}\|_{\overline{P}_0} + \|h_n \circ \nu_{n,\diamond} - h_{n,0} \circ \nu_{n,\diamond}\|_{\overline{P}_0} = O_p(n^{-1/3}),$$

and hence

$$\|h_n^{\#} \circ \nu_{n,\diamond} - h_n \circ \nu_{n,\diamond}\|_{\overline{P}_0}^2 = O_p(n^{-2/3}).$$

Therefore, for any function $\overline{\nu}_0$,

$$\begin{aligned}\|h_n^{\#} \circ \nu_{n,\diamond} - \overline{\nu}_0\|_{\overline{P}_0} &\leq \|h_n^{\#} \circ \nu_{n,\diamond} - h_n \circ \nu_{n,\diamond}\|_{\overline{P}_0} + \|h_n \circ \nu_{n,\diamond} - \overline{\nu}_0\|_{\overline{P}_0} \\ &\leq O_p(n^{-1/3}) + \|h_n \circ \nu_{n,\diamond} - \overline{\nu}_0\|_{\overline{P}_0}.\end{aligned}$$

□

The following theorem is a direct application of Theorem A.14. In the following, we define the population minimizers corresponding to the calibrators f_n and g_n as:

$$f_{n,0} := \operatorname*{argmin}_{f \in \mathcal{F}_{iso}} \sum_{j=1}^{J} \|\mu_0 - f \circ \mu_{n,j}\|_{P_0}^2;$$

$$g_{n,0} := \operatorname*{argmin}_{g \in \mathcal{F}_{iso}} \sum_{j=1}^{J} \left\{ \|g \circ \alpha_{n,j}\|_{P_0}^2 - 2\psi_0(g \circ \alpha_{n,j}) \right\}.$$

Theorem A.15. *Suppose that Conditions C5, C3, and Condition D3 hold. Then,* $\|\alpha_{n,\diamond}^* - g_{n,0} \circ \alpha_{n,\diamond}\|_{\overline{P}_0} = O_p(n^{-1/3})$ *and* $\|\mu_{n,\diamond}^* - f_{n,0} \circ \mu_{n,\diamond}\|_{\overline{P}_0} = O_p(n^{-1/3})$. *Hence,*

$$\|\alpha_{n,\diamond}^* - \alpha_0\|_{\overline{P}_0} \leq \min_{g \in \mathcal{F}_{iso}} \|g \circ \alpha_{n,\diamond} - \alpha_0\| + O_p(n^{-1/3});$$

$$\|\mu_{n,\diamond}^* - \mu_0\|_{\overline{P}_0} \leq \min_{f \in \mathcal{F}_{iso}} \|f \circ \mu_{n,\diamond} - \mu_0\| + O_p(n^{-1/3}).$$

Additionally, assume D1 holds. Then, the bootstrap calibrated nuisance estimators satisfy, for any functions $\overline{\alpha}_0, \overline{\mu}_0 \in L^2(P_0)$ *and as* $n \to \infty$, *the following root mean square error*

bounds:

$$\|\alpha^{\#}_{n,\diamond} - \overline{\alpha}_0\|_{\overline{P}_0} \leq \|\alpha^{*}_{n,\diamond} - \overline{\alpha}_0\|_{\overline{P}_0} + O_p(n^{-1/3});$$
$$\|\mu^{\#}_{n,\diamond} - \overline{\mu}_0\|_{\overline{P}_0} \leq \|\mu^{*}_{n,\diamond} - \overline{\mu}_0\|_{\overline{P}_0} + O_p(n^{-1/3}).$$

Proof. The proof follows from two direct applications of Theorem A.14. To obtain the claimed rates for $\alpha^{*}_{n,\diamond}$, we take

$$\ell(z,\nu) := \nu^2(a,w) - 2m(z,\nu), \quad h_n := g_n, \quad h_{n,0} := g_{n,0}, \quad h^{\#}_n := g^{\#}_n, \quad \nu_0 := \alpha_0, \quad \nu_{n,j} := \alpha_{n,j}.$$

Similarly, to obtain the rates for $\mu^{*}_{n,\diamond}$, we take

$$\ell(z,\nu) := \nu^2(a,w) - 2y\nu(a,w), \quad h_n := f_n, \quad h_{n,0} := f_{n,0}, \quad h^{\#}_n := f^{\#}_n, \quad \nu_0 := \mu_0, \quad \nu_{n,j} := \mu_{n,j}.$$

In both cases, Condition F1 holds by Conditions C5 and D1. For the least-squares loss, Condition F2 holds trivially, while for the Riesz loss, it holds by Condition A1. For the least-squares loss, Condition F3 holds by the pointwise Lipschitz continuity of the map $m \mapsto m^2 - 2ym$ and the uniform boundedness of the outcome Y. For the Riesz loss, Condition F3 holds by D3.

□

A.5 Proofs of main results

A.5.1 Proof of Theorem 2.4

Let $\alpha^{*}_{n,\diamond} := \{\alpha^{*}_{n,j} : j \in [J]\}$ and $\mu^{*}_{n,\diamond} := \{\mu^{*}_{n,j} : j \in [J]\}$ denote the isotonic-calibrated nuisance estimators produced by Algorithm 1.

The following lemma shows that isotonic calibration of the cross-fitted nuisance estimators ensures empirical calibration for their respective loss functions. Consequently, the calibrated nuisance estimators satisfy cross-fitted orthogonality conditions, and the leading bias terms in the expansion of the cross-product remainder in Lemma A.2 vanish.

Lemma A.16 (Empirical calibration via isotonic calibration)**.** *Assume C3. Then, for all bounded measurable $\theta_1, \theta_2 : \mathbb{R} \to \mathbb{R}$, it holds that*

$$\frac{1}{n}\sum_{i=1}^{n}(\theta_1 \circ \mu^{*}_{n,j(i)})(A_i, W_i)\left\{Y_i - \mu^{*}_{n,j(i)}(A_i, W_i)\right\} = 0;$$
$$\frac{1}{n}\sum_{i=1}^{n}\left\{(\theta_2 \circ \alpha^{*}_{n,j(i)})(A_i, W_i)\alpha^{*}_{n,j(i)}(A_i, W_i) - \psi_{n,j(i)}(\theta_2 \circ \alpha^{*}_{n,j(i)})\right\} = 0.$$

As a consequence,

$$\frac{1}{n}\sum_{i=1}^{n} s^{*}_{n,j(i)}(A_i, W_i)\left\{Y_i - \mu^{*}_{n,j(i)}(A_i, W_i)\right\} = 0;$$
$$\frac{1}{n}\sum_{i=1}^{n}\left\{r^{*}_{n,j(i)}(A_i, W_i)\alpha^{*}_{n,j(i)}(A_i, W_i) - \psi_{n,j(i)}(r^{*}_{n,j(i)})\right\} = 0.$$

Proof. The second part of the lemma follows from the first part because, by definition of the projection, each $s^{*}_{n,j}$ can be written as $s^{*}_{n,j} = \theta \circ \mu^{*}_{n,j}$ for a common map θ across $j \in [J]$,

where $s^*_{n,j} = (\overline{\Pi}_{\mu^*_{n,\diamond}}(\alpha_0 - \alpha_{n,\diamond}))_j$. Similarly, by definition of the projection, each $r^*_{n,j}$ can be written as $r^*_{n,j} = \vartheta \circ \alpha^*_{n,j}$ for a common map ϑ, where $r^*_{n,j} = (\overline{\Pi}_{\alpha^*_{n,\diamond}}(\mu_0 - \mu_{n,\diamond}))_j$.

The proof of this lemma follows from a simplification of the proof of Lemma 4 in Van Der Laan et al. [2023]. We give the proof for $\alpha^*_{n,\diamond}$, as the argument for $\mu^*_{n,\diamond}$ follows along similar lines. Recall that $\alpha^*_{n,\diamond} = g_n \circ \alpha_{n,\diamond}$, where $g_n \in \mathcal{F}_{\text{iso}}$ is the empirical risk minimizer defined in Algorithm 1. The solution g_n is unique only in $L^2(P_n)$. Following Groeneboom and Lopuhaa [1993] and without loss of generality, we take g_n to be the unique càdlàg piecewise-constant solution of the isotonic regression problem that can jump only at the observed values $\{\alpha_{n,j(i)}(A_i, W_i) : i \in [n]\}$.

By Condition B2, the empirical risk used to define the isotonic solution g_n is strongly convex and continuous on $\mathcal{F}_{\text{iso}}$ with respect to the empirical L^2 norm $\left\{\sum_{i=1}^n \left[(g \circ \alpha_{n,j(i)})(A_i, W_i)\right]^2\right\}^{1/2}$. By Theorem 5.5 of Alexanderian [2019], a strongly convex and continuous functional defined on a closed and bounded convex subset of a Hilbert space admits a unique minimizing solution on the set $\{\alpha_{n,j(i)}(A_i, W_i) : i \in [n]\}$. Consequently, the solution g_n is uniquely defined on the set $\{\alpha_{n,j(i)}(A_i, W_i) : 1 \leq i \leq n\}$, and, by assumption, it is the uniquely defined piecewise constant solution with jumps occurring only at these values.

It holds that g_n is piecewise constant. We claim that, for any function $h : \mathbb{R} \to \mathbb{R}$ and all $\varepsilon \in \mathbb{R}$ sufficiently close to zero, the perturbed function $g_n + \varepsilon h \circ g_n$ belongs to $\mathcal{F}_{iso}$. To see this, observe that g_n is constant on some interval of its domain if and only if the same holds for $h \circ g_n$. Hence, the function $g_n + \varepsilon h \circ g_n$ is piecewise constant and jumps at the same points as g_n. Moreover, by taking ε sufficiently small, we can ensure that the maximal jump change $\varepsilon \sup_{i \in [n]} |(h \circ g_n)(\mu_{n,j(i)}(A_i, W_i))|$ of $h \circ g_n$ is strictly smaller than the maximal jump change of g_n. Therefore, for this choice of ε, the perturbed function $g_n + \varepsilon h \circ g_n \in \mathcal{F}_{iso}$ must be nondecreasing and, therefore, an element of $\mathcal{F}_{iso}$. The claim now follows.

Proceeding with the proof of the lemma, we denote $\xi_n(\varepsilon) := g_n + \varepsilon(h \circ g_n)$ and $\xi_n(\varepsilon, \alpha_{n,j}) := g_n \circ \alpha_{n,j} + \varepsilon(h \circ g_n)(\alpha_{n,j})$ and, in a slight abuse of notation, let

$$R_n(\xi_n(\varepsilon)) := \sum_{i=1}^n \left\{\xi_n(\varepsilon, \alpha_{n,j(i)})^2(A_i, W_i) - 2\psi_{n,j(i)}(\xi_n(\varepsilon, \alpha_{n,j(i)}))\right\}.$$

Now, because g_n minimises $g \mapsto R_n(g)$ over $\mathcal{F}_{\text{iso}}$ and $g_n + \varepsilon h \circ g_n \in \mathcal{F}_{\text{iso}}$ for ε sufficiently close to 0, g_n also minimizes $g \mapsto R_n(g)$ over the one-dimensional submodel $\varepsilon \mapsto g_n + \varepsilon h \circ g_n$ in a neighborhood of $\varepsilon = 0$. Thus, the first-order optimality condition along this submodel implies that

$$\frac{d}{d\varepsilon}[R_n(\xi_n(\varepsilon)) - R_n(g_n)]\Big|_{\varepsilon=0} = 0.$$

Computing the derivative, using linearity of the functional $\psi_{n,\diamond}$ and that $h \circ g_n \circ \alpha_{n,\diamond} = h \circ \alpha^*_{n,\diamond}$, this further implies:

$$\frac{1}{n}\sum_{i=1}^n \left\{(h \circ \alpha^*_{n,j(i)})(A_i, W_i)\alpha^*_{n,j(i)}(A_i, W_i) - \psi_{n,j(i)}(h \circ \alpha^*_{n,j(i)})\right\} = 0.$$

The result for $\alpha^*_{n,\diamond}$ now follows, noting that h was an arbitrary bounded function.

□

Proof of Theorem 2.4. **Outline.** Our proof proceeds in three steps. First, we show that the

estimation error $\tau_n^* - \tau_0$ admits the decomposition

$$\tau_n^* - \tau_0 = (P_n - P_0)\{\phi_{0,\overline{\mu}_0} + D_{\overline{\mu}_0,\overline{\alpha}_0}\} + o_p(n^{-1/2}) \; + \; \langle \alpha_{n,\diamond}^* - \alpha_0,\; \mu_0 - \mu_{n,\diamond}^* \rangle_{\overline{P}_0},$$

that is, an asymptotically linear term plus the usual cross-product remainder. Second, we use the orthogonality properties implied by nuisance calibration to control this cross-product remainder under each of the two cases in Condition C1. Specifically, we show that

$$\langle \alpha_{n,\diamond}^* - \alpha_0,\; \mu_0 - \mu_{n,\diamond}^* \rangle_{\overline{P}_0} = 1(\overline{\alpha}_0 \neq \alpha_0)\, P_n A_0 + 1(\overline{\mu}_0 \neq \mu_0)\, P_n B_0 + o_p(n^{-1/2}).$$

Finally, combining these steps establishes that τ_n^* is DRAL with influence function χ_0.

Starting point: decomposition of the estimation error into an asymptotically linear term and a cross-product drift term. By Lemma A.1, we have the standard doubly robust expansion

$$\tau_n^* - \tau_0 = (\overline{P}_n - \overline{P}_0)\left\{\phi_{0,\mu_{n,\diamond}^*} + D_{\mu_{n,\diamond}^*,\alpha_{n,\diamond}^*}\right\} + \langle \alpha_{n,\diamond}^* - \alpha_0,\; \mu_0 - \mu_{n,\diamond}^* \rangle_{\overline{P}_0} + Rem_{\mu_{n,\diamond}^*}(\psi_{n,\diamond}, \psi_0), \tag{A.10}$$

where $Rem_{\mu_{n,\diamond}^*}(\psi_{n,\diamond}, \psi_0) := \frac{1}{J}\sum_{j=1}^J \left\{\psi_{n,j}(\mu_{n,j}^*) - \psi_0(\mu_{n,j}^*) - P_{n,j}\phi_{0,\mu_{n,j}^*}\right\}$. By the locally uniform asymptotic linearity condition in C4, we have

$$\begin{aligned} Rem_{\mu_{n,\diamond}^*}(\psi_{n,\diamond}, \psi_0) &= (\overline{P}_n - \overline{P}_0)\left\{\phi_{0,\overline{\mu}_0} - \phi_{0,\mu_{n,\diamond}^*}\right\} + \frac{1}{J}\sum_{j=1}^J \left\{\psi_{n,j}(\mu_{n,j}^*) - \psi_0(\mu_{n,j}^*) - P_{n,j}\phi_{0,\overline{\mu}_0}\right\}. \\ &= (\overline{P}_n - \overline{P}_0)\left\{\phi_{0,\overline{\mu}_0} - \phi_{0,\mu_{n,\diamond}^*}\right\} + o_p(n^{-1/2}), \end{aligned}$$

where we used that $\overline{P}_0\phi_{0,\mu_{n,\diamond}^*} = \overline{P}_0\phi_{0,\overline{\mu}_0} = 0$, as influence functions are mean zero. Substituting the preceding display into (A.10) and collecting like terms, we obtain

$$\tau_n^* - \tau_0 = (\overline{P}_n - \overline{P}_0)\left\{\phi_{0,\overline{\mu}_0} + D_{\mu_{n,\diamond}^*,\alpha_{n,\diamond}^*}\right\} + \langle \alpha_{n,\diamond}^* - \alpha_0,\; \mu_0 - \mu_{n,\diamond}^* \rangle_{\overline{P}_0} + o_p(n^{-1/2}). \tag{A.11}$$

By Lemma A.11, $(\overline{P}_n - \overline{P}_0)\left\{D_{\mu_{n,\diamond}^*,\alpha_{n,\diamond}^*} - D_{\overline{\mu}_0,\overline{\alpha}_0}\right\} = o_p(n^{-1/2})$. Hence,

$$\begin{aligned} (\overline{P}_n - \overline{P}_0)D_{\mu_{n,\diamond}^*,\alpha_{n,\diamond}^*} &= (\overline{P}_n - \overline{P}_0)D_{\overline{\mu}_0,\overline{\alpha}_0} + (\overline{P}_n - \overline{P}_0)\left\{D_{\mu_{n,\diamond}^*,\alpha_{n,\diamond}^*} - D_{\overline{\mu}_0,\overline{\alpha}_0}\right\} \\ &= (\overline{P}_n - \overline{P}_0)D_{\overline{\mu}_0,\overline{\alpha}_0} + o_p(n^{-1/2}). \end{aligned}$$

Substituting the above into (A.11), we obtain

$$\tau_n^* - \tau_0 = (P_n - P_0)\{\phi_{0,\overline{\mu}_0} + D_{\overline{\mu}_0,\overline{\alpha}_0}\} + \langle \alpha_{n,\diamond}^* - \alpha_0,\; \mu_0 - \mu_{n,\diamond}^* \rangle_{\overline{P}_0} + o_p(n^{-1/2}). \tag{A.12}$$

The above expansion is the starting point for our DRAL analysis. In what follows, we control the cross-product remainder $\langle \alpha_{n,\diamond}^* - \alpha_0,\; \mu_0 - \mu_{n,\diamond}^* \rangle_{\overline{P}_0}$ separately under the two cases in C1.

Establishing asymptotic linearity when Riesz representer is estimated well. Suppose the theorem is applied under C1i ($\overline{\alpha}_0 = \alpha_0$), so that $\|\alpha_{n,j}^* - \alpha_0\| = o_p(n^{-1/4})$ and $\|\mu_{n,j}^* - \overline{\mu}_0\| = o_p(1)$ for $j = 1, \ldots, J$. Applying the second part of Lemma A.2, we have the

Riesz-based decomposition:

$$\langle \alpha_0 - \alpha^*_{n,\diamond},\ \mu^*_{n,\diamond} - \mu_0 \rangle_{\overline{P}_0} = -\left\{\overline{\psi}_n(r^*_{n,\diamond}) - \overline{P}_n\big(r^*_{n,\diamond}\,\alpha^*_{n,\diamond}\big)\right\} + \left\langle (\overline{\Pi}_{(\alpha^*_{n,\diamond},\alpha_0)} - \overline{\Pi}_{\alpha^*_{n,\diamond}})\Delta_{\mu,n,\diamond},\ \alpha_0 - \overline{\Pi}_{\alpha^*_{n,\diamond}}\alpha_0 \right\rangle_{\overline{P}_0} + (\overline{P}_n - \overline{P}_0)B^*_{n,\diamond} + Rem_{\psi_{n,\diamond}}(r^*_{n,\diamond}), \tag{A.13}$$

where $Rem_{\psi_{n,\diamond}}(r^*_{n,\diamond}) := \frac{1}{J}\sum_{j=1}^{J}\left\{\psi_{n,j}(r^*_{n,j}) - \psi_0(r^*_{n,j}) - P_{n,j}\phi_{0,r^*_{n,j}}\right\}$. By the second part of Lemma A.12, we have

$$(\overline{P}_n - \overline{P}_0)\{B^*_{n,\diamond} - B_0\} + Rem_{\psi_{n,\diamond}}(r^*_{n,\diamond}) = o_p(n^{-1/2}).$$

Hence,

$$(\overline{P}_n - \overline{P}_0)B^*_{n,\diamond} + Rem_{\psi_{n,\diamond}}(r^*_{n,\diamond}) = (\overline{P}_n - \overline{P}_0)B_0 + o_p(n^{-1/2}).$$

Substituting the above into (A.13) and using that $\overline{P}_0 B_0 = 0$, we obtain

$$\langle \alpha_0 - \alpha^*_{n,\diamond},\ \mu^*_{n,\diamond} - \mu_0 \rangle_{\overline{P}_0} = -\left\{\overline{\psi}_n(r^*_{n,\diamond}) - \overline{P}_n\big(r^*_{n,\diamond}\,\alpha^*_{n,\diamond}\big)\right\} + \left\langle (\overline{\Pi}_{(\alpha^*_{n,\diamond},\alpha_0)} - \overline{\Pi}_{\alpha^*_{n,\diamond}})\Delta_{\mu,n,\diamond},\ \alpha_0 - \overline{\Pi}_{\alpha^*_{n,\diamond}}\alpha_0 \right\rangle_{\overline{P}_0} + \overline{P}_n B_0 + o_p(n^{-1/2}). \tag{A.14}$$

Since $\alpha^*_{n,\diamond}$ is calibrated using isotonic regression, the second part of Lemma A.16 implies that

$$-\left\{\overline{\psi}_n(r^*_{n,\diamond}) - \overline{P}_n\big(r^*_{n,\diamond}\,\alpha^*_{n,\diamond}\big)\right\} = 0.$$

Thus, the first term in (A.14) vanishes, and we conclude that

$$\langle \alpha_0 - \alpha^*_{n,\diamond},\ \mu^*_{n,\diamond} - \mu_0 \rangle_{\overline{P}_0} = \left\langle (\overline{\Pi}_{(\alpha^*_{n,\diamond},\alpha_0)} - \overline{\Pi}_{\alpha^*_{n,\diamond}})\Delta_{\mu,n,\diamond},\ \alpha_0 - \overline{\Pi}_{\alpha^*_{n,\diamond}}\alpha_0 \right\rangle_{\overline{P}_0} + \overline{P}_n B_0 + o_p(n^{-1/2}). \tag{A.15}$$

It remains to control the final term. By C2i, which holds under C1i, we have that

$$\left\langle (\overline{\Pi}_{(\alpha^*_{n,\diamond},\alpha_0)} - \overline{\Pi}_{\alpha^*_{n,\diamond}})\Delta_{\mu,n,\diamond},\ \alpha_0 - \overline{\Pi}_{\alpha^*_{n,\diamond}}\alpha_0 \right\rangle_{\overline{P}_0} = O_p\left(\|\alpha^*_{n,\diamond} - \alpha_0\|^2_{\overline{P}_0}\right).$$

Moreover, by C2i, we have $\|\alpha^*_{n,\diamond} - \alpha_0\|^2_{\overline{P}_0} = o_p(n^{-1/2})$. Hence,

$$\left\langle (\overline{\Pi}_{(\alpha^*_{n,\diamond},\alpha_0)} - \overline{\Pi}_{\alpha^*_{n,\diamond}})\Delta_{\mu,n,\diamond},\ \alpha_0 - \overline{\Pi}_{\alpha^*_{n,\diamond}}\alpha_0 \right\rangle_{\overline{P}_0} = o_p(n^{-1/2}).$$

Substituting the above into (A.15), we conclude that

$$\langle \alpha_0 - \alpha^*_{n,\diamond},\ \mu^*_{n,\diamond} - \mu_0 \rangle_{\overline{P}_0} = P_n B_0 + o_p(n^{-1/2}). \tag{A.16}$$

Substituting the above into (A.12), we obtain

$$\begin{aligned}\tau^*_n - \tau_0 &= (P_n - P_0)\left\{\phi_{0,\overline{\mu}_0} + D_{\overline{\mu}_0,\overline{\alpha}_0}\right\} + P_n B_0 + o_p(n^{-1/2}) \\ &= P_n\{B_0 + \phi_{0,\overline{\mu}_0}\} + (P_n - P_0)D_{\overline{\mu}_0,\overline{\alpha}_0} + o_p(n^{-1/2}) \\ &= P_n\{B_0 + \phi_{0,\overline{\mu}_0} + D_{\overline{\mu}_0,\overline{\alpha}_0}\} - P_0 D_{\overline{\mu}_0,\overline{\alpha}_0} + o_p(n^{-1/2}).\end{aligned}$$

By the Riesz representation property and the fact that $\overline{\alpha}_0 = \alpha_0$, we have

$$P_0 D_{\overline{\mu}_0,\overline{\alpha}_0} = P_0[\alpha_0\{\mu_0 - \overline{\mu}_0\}] = \psi_0(\mu_0 - \overline{\mu}_0).$$

Hence,

$$\begin{aligned}\tau^*_n - \tau_0 &= P_n\{B_0 + \phi_{0,\overline{\mu}_0} + D_{\overline{\mu}_0,\overline{\alpha}_0} - \psi_0(\mu_0 - \overline{\mu}_0)\} + o_p(n^{-1/2}) \\ &= P_n\chi_0 + o_p(n^{-1/2}),\end{aligned}$$

as desired.

Establishing asymptotic linearity when outcome regression is estimated well. Suppose the theorem is applied under C1ii ($\overline{\mu}_0 = \mu_0$), so that $\|\mu^*_{n,j} - \mu_0\| = o_p(n^{-1/4})$ and $\|\alpha^*_{n,j} - \overline{\alpha}_0\| = o_p(1)$ for $j = 1, \ldots, J$. Applying the first part of Lemma A.2, we have the regression-based decomposition:

$$\langle \alpha_0 - \alpha^*_{n,\diamond},\ \mu^*_{n,\diamond} - \mu_0 \rangle_{\overline{P}_0} = -\overline{P}_n A^*_{n,\diamond} + (\overline{P}_n - \overline{P}_0) A^*_{n,\diamond} + \left\langle (\overline{\Pi}_{(\mu^*_{n,\diamond},\mu_0)} - \overline{\Pi}_{\mu^*_{n,\diamond}}) \Delta_{\alpha,n,\diamond},\ \mu_0 - \overline{\Pi}_{\mu^*_{n,\diamond}} \mu_0 \right\rangle_{\overline{P}_0}. \tag{A.17}$$

By the first part of Lemma A.12, we have $(\overline{P}_n - \overline{P}_0)\{A^*_{n,\diamond} - A_0\} = o_p(n^{-1/2})$ and, hence,

$$\begin{aligned}(\overline{P}_n - \overline{P}_0) A^*_{n,\diamond} &= (\overline{P}_n - \overline{P}_0) A_0 + (\overline{P}_n - \overline{P}_0)\{A^*_{n,\diamond} - A_0\} \\ &= P_n A_0 + o_p(n^{-1/2}),\end{aligned}$$

where we used that $\overline{P}_0 A_0 = 0$. Substituting the above into (A.17), we find that

$$\begin{aligned}\langle \alpha_0 - \alpha^*_{n,\diamond},\ \mu^*_{n,\diamond} - \mu_0 \rangle_{\overline{P}_0} = &-\overline{P}_n A^*_{n,\diamond} + \overline{P}_n A_0 + o_p(n^{-1/2}) \\ &+ \left\langle (\overline{\Pi}_{(\mu^*_{n,\diamond},\mu_0)} - \overline{\Pi}_{\mu^*_{n,\diamond}}) \Delta_{\alpha,n,\diamond},\ \mu_0 - \overline{\Pi}_{\mu^*_{n,\diamond}} \mu_0 \right\rangle_{\overline{P}_0}.\end{aligned} \tag{A.18}$$

Since $\mu^*_{n,\diamond}$ is calibrated using isotonic regression, the first part of Lemma A.16 establishes that

$$\overline{P}_n A^*_{n,\diamond} = \frac{1}{n} \sum_{i=1}^{n} s^*_{n,j(i)}(A_i, W_i) \left\{ Y_i - \mu^*_{n,j(i)}(A_i, W_i) \right\} = 0.$$

Thus, (A.18) implies

$$\langle \alpha_0 - \alpha^*_{n,\diamond},\ \mu^*_{n,\diamond} - \mu_0 \rangle_{\overline{P}_0} = P_n A_0 + o_p(n^{-1/2}) + \left\langle (\overline{\Pi}_{(\mu^*_{n,\diamond},\mu_0)} - \overline{\Pi}_{\mu^*_{n,\diamond}}) \Delta_{\alpha,n,\diamond},\ \mu_0 - \overline{\Pi}_{\mu^*_{n,\diamond}} \mu_0 \right\rangle_{\overline{P}_0}. \tag{A.19}$$

It remains to control the final term. By C2ii, which holds under C1ii, we have that

$$\left\langle (\overline{\Pi}_{(\mu^*_{n,\diamond},\mu_0)} - \overline{\Pi}_{\mu^*_{n,\diamond}}) \Delta_{\alpha,n,\diamond},\ \mu_0 - \overline{\Pi}_{\mu^*_{n,\diamond}} \mu_0 \right\rangle_{\overline{P}_0} = O_p\left(\|\mu^*_{n,\diamond} - \mu_0\|^2_{\overline{P}_0} \right).$$

Moreover, by C2ii, we have $\|\mu^*_{n,\diamond} - \mu_0\|^2_{\overline{P}_0} = o_p(n^{-1/2})$. Hence,

$$\left\langle (\overline{\Pi}_{(\mu^*_{n,\diamond},\mu_0)} - \overline{\Pi}_{\mu^*_{n,\diamond}}) \Delta_{\alpha,n,\diamond},\ \mu_0 - \overline{\Pi}_{\mu^*_{n,\diamond}} \mu_0 \right\rangle_{\overline{P}_0} = o_p(n^{-1/2}).$$

Substituting the above into (A.19), we conclude that

$$\langle \alpha_0 - \alpha^*_{n,\diamond},\ \mu^*_{n,\diamond} - \mu_0 \rangle_{\overline{P}_0} = P_n A_0 + o_p(n^{-1/2}). \tag{A.20}$$

Substituting the above into (A.12), we obtain

$$\begin{aligned}\tau^*_n - \tau_0 &= (P_n - P_0)\left\{ \phi_{0,\overline{\mu}_0} + D_{\overline{\mu}_0,\overline{\alpha}_0} \right\} + P_n A_0 + o_p(n^{-1/2}) \\ &= P_n \left\{ \phi_{0,\overline{\mu}_0} + D_{\overline{\mu}_0,\overline{\alpha}_0} + A_0 \right\} + o_p(n^{-1/2}) \\ &= P_n \chi_0 + o_p(n^{-1/2}),\end{aligned}$$

where we used that $P_0 D_{\overline{\mu}_0,\overline{\alpha}_0} = 0$ when $\overline{\mu}_0 = \mu_0$. The claim follows.

Completing the proof. We have now shown that if either of the two conditions in C1 holds, then

$$\tau_n^* - \tau_0 = P_n \chi_0 + o_p(n^{-1/2}).$$

Hence, τ_n^* is DRAL. If both nuisance functions are consistently estimated ($\overline{\alpha}_0 = \alpha_0$ and $\overline{\mu}_0 = \mu_0$), then the above display reduces to

$$\tau_n^* - \tau_0 = P_n D_0 + o_p(n^{-1/2}),$$

since $\chi_0 = D_0$ in this case. By Theorem 2.1, this implies that τ_n^* is asymptotically linear with influence function equal to the efficient influence function of Ψ under the nonparametric model $\mathcal{M}$. By the characterization of efficient estimators in Van der Vaart [2000, Lemma 25.23], a regular estimator is asymptotically efficient only if it is asymptotically linear with influence function equal to the efficient influence function; see also Bickel et al. [1993]. Hence, τ_n^* is a regular and efficient estimator of τ_0 under the nonparametric model $\mathcal{M}$.

□

A.5.2 Proof of Theorem 2.5

Proof of Theorem 2.5. **Starting point.** By Lemma A.1, it holds that

$$\begin{aligned}\tau_n^\# - \tau_0 &= (\overline{P}_n^\# - \overline{P}_0)\left\{\phi_{0,\mu_{n,\diamond}^\#} + D_{\mu_{n,\diamond}^\#,\alpha_{n,\diamond}^\#}\right\} + \langle \alpha_{n,\diamond}^\# - \alpha_0,\ \mu_0 - \mu_{n,\diamond}^\# \rangle_{\overline{P}_0} + Rem_{\mu_{n,\diamond}^\#}(\psi_{n,\diamond}, \psi_0) \\ &= (\overline{P}_n^\# - \overline{P}_0)\left\{m(\cdot, \mu_{n,\diamond}^\#) + D_{\mu_{n,\diamond}^\#,\alpha_{n,\diamond}^\#}\right\} + \langle \alpha_{n,\diamond}^\# - \alpha_0,\ \mu_0 - \mu_{n,\diamond}^\# \rangle_{\overline{P}_0},\end{aligned}$$

where we used that $\phi_{0,\mu_{n,\diamond}^\#} = m(\cdot, \mu_{n,\diamond}^\#) - \overline{P}_0 m(\cdot, \mu_{n,\diamond}^\#)$ and that $Rem_{\mu_{n,\diamond}^\#}(\psi_{n,\diamond}, \psi_0) = 0$ in the setup of Section 2.5.3. Arguing as in the proof of Theorem 2.4, we may apply the second part of Lemma A.11 to obtain

$$(\overline{P}_n^\# - \overline{P}_0)\left\{m(\cdot, \mu_{n,\diamond}^\#) + D_{\mu_{n,\diamond}^\#,\alpha_{n,\diamond}^\#}\right\} = (\overline{P}_n^\# - \overline{P}_0)\left\{m(\cdot, \overline{\mu}_0) + D_{\overline{\mu}_0,\overline{\alpha}_0}\right\} + o_p(n^{-1/2}).$$

The conditions of Theorem A.15 assume the setup of Theorem 2.5, together with D3. Hence, we may apply Theorem A.15 to conclude that

$$\begin{aligned}\|\alpha_{n,\diamond}^\# - \alpha_0\|_{\overline{P}_0} &= O_p\big(\|\alpha_{n,\diamond}^* - \alpha_0\|_{\overline{P}_0}\big) + O_p(n^{-1/3}), \\ \|\mu_{n,\diamond}^\# - \mu_0\|_{\overline{P}_0} &= O_p\big(\|\mu_{n,\diamond}^* - \mu_0\|_{\overline{P}_0}\big) + O_p(n^{-1/3}).\end{aligned}$$

Furthermore, by Lemma A.10, under Conditions C5, C3, and D3, whenever one part of C1 holds for $\alpha_{n,\diamond}^*$ and $\mu_{n,\diamond}^*$, the same part holds for $\alpha_{n,\diamond}^\#$ and $\mu_{n,\diamond}^\#$. In particular, under C1, we have $\|\alpha_{n,\diamond}^\# - \overline{\alpha}_0\|_{\overline{P}_0} = o_p(1)$ and $\|\mu_{n,\diamond}^\# - \overline{\mu}_0\|_{\overline{P}_0} = o_p(1)$. Moreover, under C1ii, $\|s_{n,j}^\# - s_0\| = o_p(1)$, and under C1i, $1\{\overline{\alpha}_0 = \alpha_0\}\,\|r_{n,j}^\# - r_0\| = o_p(1)$. Consequently, under (D2), we can apply the proof of Theorem 2.4 with the original dataset $\mathcal{D}_n$ replaced by the bootstrap data $\mathcal{D}_n^\#$ and nuisance estimators replaced by their bootstrap counterparts.

We now establish asymptotic linearity of the cross-product remainder in the two cases.

Establishing asymptotic linearity when the outcome regression is estimated well. Suppose the theorem is applied under C1ii ($\overline{\mu}_0 = \mu_0$), so that $\|\mu_{n,j}^* - \mu_0\| = o_p(n^{-1/4})$ and

$\|\alpha^*_{n,j} - \overline{\alpha}_0\| = o_p(1)$ for $j = 1, \ldots, J$. The first part of Lemma A.12 implies that

$$(\overline{P}^{\#}_n - P_0)\{A^{\#}_{n,\diamond} - A_0\} = o_p(n^{-1/2}).$$

Applying the regression-based decomposition in Lemma A.2 to the bootstrap nuisance estimators, and arguing as in the proof of Theorem 2.4 using the partial orthogonality and projection error bounds in D2 together with the first part of Lemma A.12, we obtain

$$\langle \alpha^{\#}_{n,\diamond} - \alpha_0,\ \mu_0 - \mu^{\#}_{n,\diamond} \rangle_{\overline{P}_0} = 1(\overline{\alpha}_0 \neq \alpha_0)\,(\overline{P}^{\#}_n - P_0)A_0 + O_p\Big(\|\mu^{\#}_{n,\diamond} - \mu_0\|^2_{\overline{P}_0}\Big).$$

By Theorem A.15, we have $\|\mu^{\#}_{n,\diamond} - \mu_0\|^2_{\overline{P}_0} = O_p(n^{-2/3}) + \|\mu^*_{n,\diamond} - \mu_0\|^2_{\overline{P}_0}$, and thus, by C1, it follows that $\|\mu^{\#}_{n,\diamond} - \mu_0\|^2_{\overline{P}_0} = o_p(n^{-1/2})$. We conclude that

$$\langle \alpha^{\#}_{n,\diamond} - \alpha_0,\ \mu_0 - \mu^{\#}_{n,\diamond} \rangle_{\overline{P}_0} = 1(\overline{\alpha}_0 \neq \alpha_0)\,(\overline{P}^{\#}_n - P_0)A_0 + o_p(n^{-1/2}).$$

Establishing asymptotic linearity when the Riesz representer is estimated well. Suppose the theorem is applied under C1i ($\overline{\alpha}_0 = \alpha_0$), so that $\|\alpha^*_{n,j} - \alpha_0\| = o_p(n^{-1/4})$ and $\|\mu^*_{n,j} - \overline{\mu}_0\| = o_p(1)$ for $j = 1, \ldots, J$. By the second part of Lemma A.12 applied in the bootstrap setting, we have

$$(\overline{P}^{\#}_n - P_0)\{B^{\#}_{n,\diamond} - B_0\} + \frac{1}{J}\sum_{j=1}^{J}\left\{\psi_{n,j}(r^{\#}_{n,j}) - \psi_0(r^{\#}_{n,j}) - P_{n,j}\phi_{0,r^{\#}_{n,j}}\right\} = o_p(n^{-1/2}).$$

Applying the Riesz-based decomposition in Lemma A.2 to the bootstrap nuisance estimators, and arguing as in the proof of Theorem 2.4 using the partial orthogonality and projection error bounds in D2 together with the second part of Lemma A.12, we obtain

$$\langle \alpha^{\#}_{n,\diamond} - \alpha_0, \mu_0 - \mu^{\#}_{n,\diamond} \rangle_{\overline{P}_0} = 1(\overline{\mu}_0 \neq \mu_0)(\overline{P}^{\#}_n - P_0)B_0 + O_p\left(\|\alpha^{\#}_{n,\diamond} - \alpha_0\|^2_{\overline{P}_0}\right).$$

By Theorem A.15, it holds that

$$\|\alpha^{\#}_{n,\diamond} - \alpha_0\|^2_{\overline{P}_0} = O_p(n^{-2/3}) + \|\alpha^*_{n,\diamond} - \alpha_0\|^2_{\overline{P}_0},$$

and, thus, by C1, $\|\alpha^{\#}_{n,\diamond} - \alpha_0\|^2_{\overline{P}_0} = o_p(n^{-1/2})$. We conclude that

$$\langle \alpha^{\#}_{n,\diamond} - \alpha_0, \mu_0 - \mu^{\#}_{n,\diamond} \rangle_{\overline{P}_0} = 1(\overline{\mu}_0 \neq \mu_0)(\overline{P}^{\#}_n - P_0)B_0 + o_p(n^{-1/2}).$$

Establishing sample-conditional DRAL of $\tau^{\#}_n - \tau^*_n$. Combining the two cases treated above, we obtain

$$\langle \alpha^{\#}_{n,\diamond} - \alpha_0,\ \mu_0 - \mu^{\#}_{n,\diamond} \rangle_{\overline{P}_0} = 1(\overline{\alpha}_0 \neq \alpha_0)\,(\overline{P}^{\#}_n - P_0)A_0 + 1(\overline{\mu}_0 \neq \mu_0)\,(\overline{P}^{\#}_n - P_0)B_0 + o_p(n^{-1/2}).$$

Hence,

$$\tau^{\#}_n - \tau_0 = \overline{P}^{\#}_n \chi_0 + o_p(n^{-1/2}).$$

By Theorem 2.4, we also have

$$\tau^*_n - \tau_0 = (\overline{P}_n - \overline{P}_0)\chi_0 + o_p(n^{-1/2}),$$

and therefore

$$\tau^{\#}_n - \tau^*_n = (\overline{P}^{\#}_n - \overline{P}_n)\chi_0 + o_p(n^{-1/2}).$$

Establishing weak convergence of bootstrap law. Let n_j denote the number of observations in $\mathcal{C}^{(j)}$ (so that $P_{n,j}$ is the empirical measure based on n_j samples). By assumption, the splits are approximately equal, in the sense that $Jn_j/n \to 1$. For each $j \in [J]$, the bootstrap CLT [Van Der Vaart and Wellner, 1996] yields

$$\sqrt{n_j}\,(P_{n,j}^{\#} - P_{n,j})\chi_0 \,\big|\, P_{n,j} \;\rightsquigarrow\; N\big(0,\, P_0\chi_0^2\big)\,.$$

Since the bootstrap resamples are generated independently across folds conditional on the data, the same limit holds conditional on the full collection $\{P_{n,j} : j \in [J]\}$:

$$\sqrt{n_j}\,(P_{n,j}^{\#} - P_{n,j})\chi_0 \,\big|\, \{P_{n,j} : j \in [J]\} \;\rightsquigarrow\; N\big(0,\, P_0\chi_0^2\big)\,.$$

Moreover,

$$(\overline{P}_n^{\#} - \overline{P}_n)\chi_0 = \frac{1}{J}\sum_{j=1}^{J}(P_{n,j}^{\#} - P_{n,j})\chi_0,$$

and $Jn_j/n \to 1$ implies $\sqrt{n}/J = \sqrt{n_j}\,(1 + o(1))$. Therefore, by Slutsky's lemma,

$$\sqrt{n}\,(\overline{P}_n^{\#} - \overline{P}_n)\chi_0 \,\big|\, \{P_{n,j} : j \in [J]\} \;\rightsquigarrow\; N\big(0,\, P_0\chi_0^2\big)\,.$$

Consequently, the conditional law of $\sqrt{n}\,(\tau_n^{\#} - \tau_n^*)$ given $\{P_{n,j} : j \in [J]\}$ converges in probability to the law of $N(0, P_0\chi_0^2)$. Indeed, writing

$$\sqrt{n}\,(\tau_n^{\#} - \tau_n^*) = \sqrt{n}\,(\overline{P}_n^{\#} - \overline{P}_n)\chi_0 + \sqrt{n}\,R_n,$$

our preceding argument shows that

$$\mathcal{L}\Big(\sqrt{n}\,(\overline{P}_n^{\#} - \overline{P}_n)\chi_0 \,\Big|\, \{P_{n,j} : j \in [J]\}\Big) \Rightarrow N(0, P_0\chi_0^2) \quad \text{in } P_0\text{-probability.}$$

Moreover, since $\sqrt{n}\,R_n = o_P(1)$ under the joint law of $(P_n, P_n^{\#})$, we have $P(|\sqrt{n}\,R_n| > \eta) \to 0$ for every $\eta > 0$. Fix $\eta, \delta > 0$ and define

$$Q_n \;:=\; P^{\#}\big(|\sqrt{n}\,R_n| > \eta \,\big|\, \{P_{n,j} : j \in [J]\}\big) \in [0,1].$$

By the tower property,

$$E_0[Q_n] \;=\; P\big(|\sqrt{n}\,R_n| > \eta\big)\,,$$

where E_0 denotes expectation with respect to the data-generating law and P denotes the joint law of $(P_n, P_n^{\#})$. Therefore, by Markov's inequality,

$$P_0(Q_n > \delta) \;\le\; \frac{E_0[Q_n]}{\delta} \;=\; \frac{P(|\sqrt{n}\,R_n| > \eta)}{\delta} \;\to\; 0.$$

Thus, $\sqrt{n}\,R_n = o_{P^{\#}}(1)$ in P_0-probability, in the sense that for every $\eta, \delta > 0$,

$$P_0\big(\Pr\big(|\sqrt{n}\,R_n| > \eta \,\big|\, \{P_{n,j} : j \in [J]\}\big) > \delta\big) \to 0.$$

By Slutsky's lemma for conditional laws, it follows that

$$\mathcal{L}\Big(\sqrt{n}\,(\tau_n^{\#} - \tau_n^*) \,\Big|\, \{P_{n,j} : j \in [J]\}\Big) \Rightarrow N(0, P_0\chi_0^2) \quad \text{in } P_0\text{-probability.}$$

That is, the law $\mathcal{L}(\sqrt{n}\,(\tau_n^{\#} - \tau_n^*) \mid \{P_{n,j} : j \in [J]\})$ converges weakly in probability to the law of $N(0, P_0\chi_0^2)$, in the sense that for every bounded Lipschitz function $f : \mathbb{R} \to \mathbb{R}$,

$$E\Big[f\Big(\sqrt{n}\,(\tau_n^{\#} - \tau_n^*)\Big) \,\Big|\, \{P_{n,j} : j \in [J]\}\Big] \;\xrightarrow{P_0}\; E[f(Z)]\,, \qquad Z \sim N(0, P_0\chi_0^2).$$

□

A.5.3 Proof of Theorem 2.13

Proof of Theorem 2.13. Recall that $\overline{\Pi}_{\alpha_0,P}$ denotes the $L^2(P)$-projection onto $\overline{\Theta}_{\alpha_0,P}$, and $\overline{\Pi}_{\mu_0,P}$ denotes the $L^2(P)$-projection onto $\overline{\Theta}_{\mu_0,P}$.

First suppose that $\overline{\alpha}_0 = \alpha_0$ and $\overline{\mu}_0 \neq \mu_0$. Then

$$\begin{aligned}\chi_0 &= m(\cdot, \overline{\mu}_0) - \psi_0(\mu_0) + D_{\overline{\mu}_0,\alpha_0} + B_0 \\ &= m(\cdot, \overline{\mu}_0) - \psi_0(\mu_0) + \alpha_0\{\mathcal{I}_Y - \overline{\mu}_0\} + m(\cdot, r_0) - r_0\alpha_0 \\ &= m(\cdot, \overline{\mu}_0 + r_0) + \alpha_0\{\mathcal{I}_Y - \overline{\mu}_0 - r_0\} - \psi_0(\mu_0).\end{aligned}$$

By the orthogonality conditions defining r_0, $\psi_0(\mu_0) = \psi_0(\overline{\mu}_0 + r_0)$. Therefore, since

$$\overline{\mu}_0 + r_0 = \overline{\Pi}_{\alpha_0,P_0}\mu_0,$$

we have

$$\begin{aligned}\chi_0 &= m(\cdot, \overline{\mu}_0 + r_0) - \psi_0(\overline{\mu}_0 + r_0) + \alpha_0\{\mathcal{I}_Y - \overline{\mu}_0 - r_0\} \\ &= m(\cdot, \overline{\Pi}_{\alpha_0,P_0}\mu_0) - \psi_0(\overline{\Pi}_{\alpha_0,P_0}\mu_0) + \alpha_0\{\mathcal{I}_Y - \overline{\Pi}_{\alpha_0,P_0}\mu_0\}.\end{aligned}$$

The right-hand side is the P_0-efficient influence function of the parameter

$$P \mapsto \psi_P(\overline{\Pi}_{\alpha_0,P}\mu_P).$$

Indeed, this is the same efficient influence function calculation as in Theorem 2.1, with the outcome Y replaced by the residual $Y - \overline{\mu}_0$. The required Hellinger-continuity and pathwise differentiability conditions remain valid because $P \mapsto \psi_P$ satisfies these conditions by assumption and the P-indexed projection map

$$P \mapsto \overline{\Pi}_{\alpha_0,P}\mu_P$$

is Hellinger continuous and pathwise differentiable at P_0 [Luedtke and Chung, 2024, Luedtke, 2024]. Hence the functional $P \mapsto \psi_P(\overline{\Pi}_{\alpha_0,P}\mu_P)$ satisfies the same chain-rule conditions used in Theorem 2.1.

Next suppose that $\overline{\alpha}_0 \neq \alpha_0$ and $\overline{\mu}_0 = \mu_0$. Then

$$\begin{aligned}\chi_0 &= m(\cdot, \mu_0) - \psi_0(\mu_0) + D_{\mu_0,\overline{\alpha}_0} + A_0 \\ &= m(\cdot, \mu_0) - \psi_0(\mu_0) + D_{\mu_0,\overline{\alpha}_0} + s_0\{\mathcal{I}_Y - \mu_0\} \\ &= m(\cdot, \mu_0) - \psi_0(\mu_0) + \{\overline{\alpha}_0 + s_0\}\{\mathcal{I}_Y - \mu_0\}.\end{aligned}$$

Since

$$\overline{\alpha}_0 + s_0 = \overline{\Pi}_{\mu_0,P_0}\alpha_0,$$

this gives

$$\chi_0 = m(\cdot, \mu_0) - \psi_0(\mu_0) + \overline{\Pi}_{\mu_0,P_0}\alpha_0\{\mathcal{I}_Y - \mu_0\}.$$

We claim that the right-hand side is the P_0-efficient influence function of the parameter

$$P \mapsto \overline{\psi}_{\mu_0,P}(\mu_P), \qquad \overline{\psi}_{\mu_0,P}(h) := \langle \overline{\Pi}_{\mu_0,P}\alpha_P, h\rangle_P.$$

Indeed, at P_0, the linear functional $h \mapsto \overline{\psi}_{\mu_0,P_0}(h)$ has Riesz representer

$$\overline{\Pi}_{\mu_0,P_0}\alpha_0 = \overline{\alpha}_0 + s_0$$

by the definition of s_0.

It remains only to check that the oracle functional satisfies the regularity conditions of Theorem 2.1. Let

$$\alpha_{P,\mathrm{proj}} := \overline{\Pi}_{\mu_0,P}\alpha_P, \qquad \alpha_{0,\mathrm{proj}} := \overline{\Pi}_{\mu_0,P_0}\alpha_0 = \overline{\alpha}_0 + s_0.$$

By Hellinger continuity of $P \mapsto \psi_P$ and continuity of the $L^2(P)$-projection onto $\overline{\Theta}_{\mu_0,P}$,

$$\|\alpha_{P,\mathrm{proj}} - \alpha_{0,\mathrm{proj}}\|_{P_0} \to 0 \qquad \text{as } P \to P_0$$

in Hellinger distance. Hence

$$\|\overline{\psi}_{\mu_0,P} - \overline{\psi}_{\mu_0,P_0}\|_{\mathrm{op}} = \|\alpha_{P,\mathrm{proj}} - \alpha_{0,\mathrm{proj}}\|_{P_0} \to 0,$$

so $P \mapsto \overline{\psi}_{\mu_0,P}$ is Hellinger continuous at P_0.

For pathwise differentiability, for each fixed $h \in \mathcal{H}$,

$$\overline{\psi}_{\mu_0,P}(h) = \langle \alpha_{P,\mathrm{proj}}, h\rangle_P.$$

Thus the same chain-rule argument used for $\psi_P(h) = \langle \alpha_P, h\rangle_P$ applies with α_P replaced by $\alpha_{P,\mathrm{proj}}$. Therefore Theorem 2.1 applies to $\overline{\psi}_{\mu_0,P}$, and its efficient influence function at P_0 is obtained by replacing α_0 with

$$\alpha_{0,\mathrm{proj}} = \overline{\Pi}_{\mu_0,P_0}\alpha_0 = \overline{\alpha}_0 + s_0.$$

This gives exactly the displayed expression for χ_0.

Combining the preceding displays with Theorem 2.4 shows that τ_n^* is P_0-asymptotically linear with influence function χ_0, which is the P_0-efficient influence function of the oracle parameter Ψ_0. Therefore, by the standard characterization of efficient estimators [Van der Vaart, 2000, Lemma 25.23], τ_n^* is a regular and efficient estimator of Ψ_0 at P_0 under the nonparametric model $\mathcal{M}$; see also Bickel et al. [1993], Van Der Vaart and Wellner [1996]. The stated limiting distribution follows directly from the definition of regularity. □

Appendix B

PROOFS FOR CHAPTER 2

B.1 EP-learner uniform consistency with fixed K-nearest neighbors

Proof of Theorem 3.6. Fix $w \in (0,1)^d$ and let $B_{K_n,n}(w)$ be the smallest closed ball centered at w containing $\{W_i : i \in \mathcal{I}_{K_n,n}(w)\}$. Work on the probability-1 event that the set $\mathcal{I}_{K_n,n}(w)$ is uniquely defined (i.e., no ties). On this event, $W_i \in B_{K_n,n}(w)$ if and only if $i \in \mathcal{I}_{K_n,n}(w)$. Then

$$\begin{aligned}
\left|\theta_{n,EP}(w) - \theta_0^-(w)\right| &= \left|\frac{1}{K_n}\sum_{i=1}^n 1\{W_i \in B_{K_n,n}(w)\}\Big[\mu_n^*(1,W_i) - \mu_n^*(0,W_i) - \{\mu_0(1,w) - \mu_0(0,w)\}\Big]\right| \\
&\le \left|\frac{1}{K_n}\sum_{i=1}^n 1\{W_i \in B_{K_n,n}(w)\}\Big[\mu_n^*(1,W_i) - \mu_n^*(0,W_i) - \{\mu_0(1,W_i) - \mu_0(0,W_i)\}\Big]\right| \\
&\quad + \left|\frac{1}{K_n}\sum_{i=1}^n 1\{W_i \in B_{K_n,n}(w)\}\Big[\mu_0(1,W_i) - \mu_0(0,W_i) - \{\mu_0(1,w) - \mu_0(0,w)\}\Big]\right| \\
&\lesssim \|\theta_{n,T}^* - \theta_0^-\|_\infty + \sup_{w' \in (0,1)^d}\sup_{x \in B_{K_n,n}(w')} L\|x - w'\|_2,
\end{aligned}$$

where L is the Lipschitz constant of μ_0. Taking a supremum over $w \in (0,1)^d$ yields

$$\sup_{w\in(0,1)^d}\left|\theta_{n,EP}(w) - \theta_0^-(w)\right| \lesssim \|\theta_{n,T}^* - \theta_0^-\|_\infty + \sup_{w' \in (0,1)^d}\sup_{x \in B_{K_n,n}(w')} L\|x - w'\|_2.$$

By Corollary 1.5 of Chenavier et al. [2022], the maximal volume of $B_{K_n,n}(w')$ over $w' \in (0,1)^d$ is $O_p(K_n \log n/n)$. This implies that the maximal radius of $B_{K_n,n}(w')$ is of order $O_p((K_n \log n/n)^{1/d})$ up to a constant depending on d. Therefore the bound simplifies to

$$\sup_{w\in(0,1)^d}\left|\theta_{n,EP}(w) - \theta_0^-(w)\right| \lesssim \|\theta_{n,T}^* - \theta_0^-\|_\infty + \mathcal{O}_p\big((K_n \log n/n)^{1/d}\big),$$

which proves the theorem. □

B.2 Preliminaries for proofs

B.2.1 Notation and conventions for proofs

For two quantities x and y, we use the expression $x \lesssim y$ to mean that x is upper bounded by y times a universal constant that may only depend on global constants, including η, M, and C, that appear the conditions of Theorems 3.2-3.5.

For a function class $\mathcal{F}$, we denote its norms by $\|\mathcal{F}\| := \sup_{f\in\mathcal{F}}\|f\|$ and $\|\mathcal{F}\|_\infty := \sup_{f\in\mathcal{G}}\|f\|_\infty$. For a uniformly bounded function class $\mathcal{F}$ and distribution P, let $N(\epsilon, \mathcal{F}, L_2(P))$ denote the $\varepsilon-$covering number [van der Vaart and Wellner, 1996] of $\mathcal{F}$

with respect to the $L_2(P)$-metric. We define the uniform entropy integral of $\mathcal{F}$ by

$$\mathcal{J}(\delta, \mathcal{F}) := \int_0^\delta \sup_Q \sqrt{\log N(\varepsilon, \mathcal{F}, L_2(Q))}\, d\varepsilon \ ,$$

where the supremum is taken over all discrete probability distributions Q.

In our proofs, we use the following notation. We let $\mathbb{I}_{g_1,g_2}$ be the indicator that takes the value 0 if both g_1 and g_0 are the identity function. We denote the nuisance rates:

$$r_n^* := n^{-\beta/(2\beta+1)} + \sqrt{k(n)\log n/n};$$
$$s_n^* := r_n^* + n^{-\gamma/(2\gamma+1)}.$$

We denote, for each $P \in \mathcal{M}$, the Neyman orthogonal loss function corresponding with R_P by $L_{\pi_P,\mu_P} := L_{\mu_P} + \Delta_{\pi_P,\mu_P}$. For $\phi \in L^2(P_{0,W})$, $j \in [J]$, and $m \in \{1,2\}$, we define pointwise, for $o \in \mathcal{O}$, the following variants of the debiasing term:

$$\overline{\Delta}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(o\,,\phi) := \frac{1}{\pi_{n,j}(a\,|\,w)} \left\{ H_{m,\mu^*_{n,j}}(a,w)\cdot\phi(w) \right\} \left\{ y - \mu^*_{n,j}(a,w) \right\};$$
$$\overline{\Delta}^{*(m)}_{\pi_{n,j},\mu_{n,j}}(o\,,\phi) := \frac{1}{\pi_{n,j}(a\,|\,w)} \left\{ H_{m,\mu_{n,j}}(a,w)\cdot\phi(w) \right\} \left\{ y - \mu^*_{n,j}(a,w) \right\}.$$

The overline encodes dependence on the transformed action-space $\mathcal{H}$ as opposed to the original action-space $\mathcal{F}$. The superscript by $*$ encodes that the residual in the definition corresponds to the debiased outcome regression estimator.

Recall $\{\mathcal{D}_n^j : j \in [J]\}$ is the partitioning of $\mathcal{D}_n := \{O_i : i \in [n]\}$ used for cross-fitting the initial nuisance estimators, $\{\mu_{n,j}, \pi_{n,j} : j \in [J]\}$, as in Algorithm 6. We define the j-th training set $\mathcal{T}_n^j := \mathcal{D}_n \backslash \mathcal{D}_n^j$. We let $P_{n,j}$ denote the empirical distribution of the jth data-fold $\mathcal{D}_n^j$ and define the fold-specific empirical process operator $f \mapsto P_{n,j}f := \int f(o) dP_{n,j}(o)$. For ease of presentation, we will use the following cross-fitting notation. For a collection of fold-specific functions $\nu_{n,\diamond} := \{\nu_{n,j} : j \in [J]\}$ that may depend on n, we define the fold-averaged empirical process operators:

$$\overline{P}_0 \nu_{n,\diamond} := \frac{1}{J}\sum_{j=1}^J P_0 \nu_{n,j};$$
$$\overline{P}_n \nu_{n,\diamond} := \frac{1}{J}\sum_{j=1}^J P_{n,j} \nu_{n,j}.$$

We define the fold-averaged norm of a collection of fold-specific functions $\nu_\diamond := \{\nu_j : j \in [J]\}$ as $\|\nu_\diamond\|_{\overline{P}_0} := \sqrt{\frac{1}{J}\sum_{j=1}^J \|\nu_j\|^2_{P_0}}$. For a function $h : \mathbb{R} \to \mathbb{R}$, we also define the composition $h \circ \nu_\diamond := \{h \circ \nu_j : j \in [J]\}$. The above notation lets us treat the collection of debiased cross-fitted estimators $\mu^*_{n,\diamond} := \{\mu^*_{n,j} : j \in [J]\}$ as extended functions defined on $\mathcal{A} \times \mathcal{W} \times [J]$. For example, in our proofs, $\overline{P}_n \overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot,\phi)$ is notation for $J^{-1}\sum_{j\in[J]} \int \overline{\Delta}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(o,\phi) dP_{n,j}(o)$.

In our proofs, for some sufficiently large $M > 0$, we will work on the following events:

(E1) $\max_{j\in[J]} \|\mu_{n,j} - \mu_0\| \le M n^{-\beta/(2\beta+1)}$

(E2) $\max_{j\in[J]} \|\pi_{n,j} - \pi_0\| \le M n^{-\gamma/(2\gamma+1)}$

(E3) $\max_{j\in[J]} \left\| \mu^*_{n,j} - \mu_0 \right\| \le M n^{-\beta/(2\beta+1)} + M\sqrt{k(n)\log n/n}.$

(E4) $\|\beta_n\|_\infty \leq M$.

Let A_n be the union of events E1-E4, and let $\mathbb{I}_{A_n}$ denote the indicator that event A_n occurs. In our proofs, we will choose $M > 0$ so that the event A_n occurs asymptotically with probability at least $1 - \varepsilon$. We claim, under C2 and C1b, that we can always choose such an $M > 0$. To see this, note, under condition C2, we can take M large enough so that events E1-E3 occur asymptotically with probability arbitrarily close to one. Under Condition C1b, we claim that Event E4 occurs with probability tending to one. To show this, recall by definition that $\mu^*_{n,j(i)} = g^{-1}(g(\mu_{n,j(i)}) + \widehat{\varphi}^T_{k(n),i}\beta_n)$. By C1b, we have that $\mu_{n,j(i)}$ and $\mu^*_{n,j(i)}$ are bounded with probability tending to one and, since the feature mapping $\varphi_{k(n)}$ is bounded by definition, $\widehat{\varphi}_{k(n),i}$ is bounded with probability tending to one. Since g is a continuous invertible function, we have that β_n must also be bounded with probability tending to one. The claim then follows taking $M > 0$ large enough so that $\|\beta_n\|_\infty \leq M$ with probability tending to one.

B.2.2 Supporting Lemmas and empirical process bounds

In this section, we present some technical lemmas used to bound empirical process remainders and second-order bias terms that appear in our proofs.

The following lemma due to Belloni et al. [2015] bounds the sup-norm sieve approximation error of the $L^2(P_0)$-projection $\Pi_{k(n)}$. For the following lemma, we define the rate $\rho_{n,\infty} := \{\log k(n)\}^\nu k(n)^{-\rho}$.

Lemma B.1. *Under C3a and C3b, it holds that*

$$\sup_{\phi\in\mathcal{H}} \|\phi - \Pi_{k(n)}\phi\|_\infty \lesssim \{\log k(n)\}^\nu \sup_{\phi\in\mathcal{H}} \inf_{f\in\mathcal{H}_{k(n)}} \|\phi - f\|_\infty \lesssim \rho_{n,\infty}.$$

Proof. Condition C3a and Proposition 3.2. of Belloni et al. [2015] imply that $\sup_{\phi\in\mathcal{H}} \|\phi - \Pi_{k(n)}\phi\|_\infty \lesssim (\log k(n))^\nu \sup_{\phi\in\mathcal{H}} \inf_{f\in\mathcal{H}_{k(n)}} \|\phi - f\|_\infty$. The second bound then follows from C3b, noting that $\sup_{\phi\in\mathcal{H}} \inf_{f\in\mathcal{H}_{k(n)}} \|\phi - f\|_\infty \lesssim k(n)^{-\rho}$. □

Lemma B.2. *Let $\mathcal{F}_1, \ldots, \mathcal{F}_k$ be given uniformly bounded function classes with $\mathcal{J}_\infty(1, \mathcal{F}_j) < \infty$ for each $j \in [k]$. Let $\varphi : \mathbb{R}^k \to \mathbb{R}$ by a Lipschitz-continuous map. Then, the function class $\mathcal{G} := \{\varphi(f_1, \ldots, f_k) : (f_1, \ldots, f_k) \in \mathcal{F}_1 \times \cdots \times \mathcal{F}_k\}$ satisfies $\log N_\infty(\varepsilon, \mathcal{G}) \lesssim \sum_{j=1}^k \log N_\infty(\varepsilon, \mathcal{F}_j)$ and, hence, $\mathcal{J}_\infty(\delta, \mathcal{G}) \lesssim \sum_{j=1}^k \mathcal{J}_\infty(\delta, \mathcal{F}_j)$.*

Proof. This result follows from the proof of Theorem 2.10.20 in van der Vaart and Wellner [1996]. □

The following lemma relates the sup-norm entropy integral of $\mathcal{H}$ to those of the function classes $\mathcal{H} - \Pi_k\mathcal{H} := \{\phi - \Pi_k\phi : \phi \in \mathcal{H}\}$ and $\Pi_k\mathcal{H} := \{\Pi_k\phi : \phi \in \mathcal{H}\}$.

Lemma B.3. *Under C3 and C4, we have, for all $\delta > 0$, that $\mathcal{J}_\infty(\delta, \mathcal{H}) \lesssim \mathcal{J}_\infty(\delta, \mathcal{F})$, $\mathcal{J}_\infty(\delta, \Pi_k\mathcal{H}) \lesssim \mathcal{J}_\infty(\{\log k(n)\}^\nu\delta, \mathcal{F})$, and $\mathcal{J}_\infty(\delta, \mathcal{H} - \Pi_k\mathcal{H}) \lesssim \mathcal{J}_\infty(\{\log k(n)\}^\nu\delta, \mathcal{F})$*

Proof. Recall, by definition, that $\mathcal{H} := \{h_i \circ \theta : \theta \in \mathcal{F}, i \in \{1,2\}\}$ where h_1 and h_2 are Lipschitz-continuous functions. Thus, by Lemma B.2, we have $\mathcal{J}_\infty(\delta, \mathcal{H}) \lesssim \mathcal{J}_\infty(\delta, \mathcal{F})$. Now, for $i \in \{1,2\}$, let $(\phi_i - \Pi_k \phi_i)$ be an element of $\mathcal{H} - \Pi_k \mathcal{H}$. Using linearity of the $L^2(P_0)$-projection and the Lebesgue constant bound of C3a, we have

$$\begin{aligned}\|(\Pi_k \phi_1) - \Pi_k \phi_2\|_\infty &= \|\Pi_k(\phi_1 - \phi_2)\|_\infty \\ &\lesssim \{\log k(n)\}^\nu \|\phi_1 - \phi_2\|_\infty.\end{aligned}$$

Thus, $\Pi_k \mathcal{H} = \{\Pi_k \phi : \phi \in \mathcal{H}\}$ is, relative to the sup-norm, a Lipschitz transformation of the function class $\mathcal{H}$ with Lipschitz constant equal to $C\{\log k(n)\}^\nu$ for some constant $C > 0$. Hence, we have $\log N_\infty(\varepsilon, \mathcal{H}_k) \lesssim \log N_\infty(\{\log k(n)\}^\nu \varepsilon, \mathcal{H})$ and, therefore, by Lemma B.2, $\log N_\infty(\varepsilon, \mathcal{H}_k) \lesssim \log N_\infty(\{\log k(n)\}^\nu \varepsilon, \mathcal{F})$. It then follows, from a change of variables, that

$$\begin{aligned}\mathcal{J}_\infty(\delta, \Pi_k \mathcal{H}) &= \int_0^\delta \sqrt{\log N_\infty(\varepsilon, \mathcal{H}_k)} d\varepsilon \\ &\lesssim \int_0^\delta \sqrt{\log N_\infty(\{\log k(n)\}^\nu \varepsilon, \mathcal{F})} d\varepsilon \\ &= \{\log k(n)\}^{-\nu} \int_0^{\{\log k(n)\}^\nu \delta} \sqrt{\log N_\infty(\varepsilon, \mathcal{F})} d\varepsilon \\ &= \{\log k(n)\}^{-\nu} \mathcal{J}_\infty(\{\log k(n)\}^\nu \delta, \mathcal{F}).\end{aligned}$$

Since $\{\log k(n)\}^{-\nu} = O(1)$ as $\nu \geq 0$, we have $\mathcal{J}_\infty(\delta, \Pi_k \mathcal{H}) \lesssim \mathcal{J}_\infty(\{\log k(n)\}^\nu \delta, \mathcal{F})$, as desired. Finally, observe that $\mathcal{H} - \mathcal{H}_k$ is a Lipschitz transformation of $\mathcal{H}$ and $\mathcal{H}_k$. Thus, by Lemma B.2, we have $\mathcal{J}_\infty(\delta, \mathcal{H} - \Pi_k \mathcal{H}) \lesssim \mathcal{J}_\infty(\delta, \mathcal{H}) + \mathcal{J}_\infty(\delta, \Pi_k \mathcal{H}) \lesssim \mathcal{J}_\infty(\{\log k(n)\}^\nu \delta, \mathcal{F})$, where the final inequality follows from our earlier claims.

□

The next lemmas provide local maximal inequalities for empirical processes in terms of the sup-norm metric entropy integral. The following lemma is a restatement of Theorem 2.1. in van de Geer [2014].

Lemma B.4. *Suppose* $\mathcal{J}_\infty(\infty, \mathcal{F}) < \infty$. *Then,* $\sqrt{E\|\mathcal{F}\|_{P_n}^2} \lesssim \|\mathcal{F}\| + \frac{\mathcal{J}_\infty(\|\mathcal{F}\|_\infty, \mathcal{F})}{\sqrt{n}}$ *and, as such,*

$$E\left[\sup_{f \in \mathcal{F}} |(P_n - P)f|\right] \lesssim n^{-1/2} \mathcal{J}_\infty(\delta, \mathcal{F}),$$

for any $\delta \geq \|\mathcal{F}\| + n^{-1/2}$.

The following lemmas provide local maximal inequalities for function classes obtained by taking the pointwise product of elements of two function classes. The first lemma follows from a modification of the proof of Theorem 2.1 in van der Vaart and Wellner [2011] where, in Equation (2.2) of their proof, we argue as in the proof of Theorem 3.1 of van de Geer [2014]. In the following, let $\mathcal{H}$ and $\mathcal{G}$ be two uniformly bounded function classes. Denote the class obtained by taking pointwise products of their respective elements as $\mathcal{HG} := \{hg : h \in \mathcal{H}, g \in \mathcal{G}\}$.

Lemma B.5. *Suppose* $\mathcal{J}_\infty(1, \mathcal{H}) < \infty$ *and* $\mathcal{J}(1, \mathcal{G}) < \infty$. *Then,*

$$E\|\mathbb{G}_n\|_{\mathcal{GH}} \lesssim \sqrt{E\|\mathcal{G}\|_{P_n}^2} \mathcal{J}_\infty\left(\sqrt{E\|\mathcal{HG}\|_{P_n}^2} / \sqrt{E\|\mathcal{G}\|_{P_n}^2}, \mathcal{H}\right) + \|\mathcal{H}\|_\infty \mathcal{J}\left(\sqrt{E\|\mathcal{HG}\|_{P_n}^2} / \|\mathcal{H}\|_\infty, \mathcal{G}\right).$$

Furthermore, $\sqrt{E\left\|\mathcal{H}\mathcal{G}\right\|_{P_n}^2} \lesssim \left\|\mathcal{G}\right\|_\infty \max\left(\left\|\mathcal{H}\right\|_{P_0}, \frac{\mathcal{J}_\infty(\left\|\mathcal{H}\right\|_\infty, \mathcal{H})}{\sqrt{n}}\right)$ *and, for any* $\delta_{crit,\mathcal{G}} > 0$ *satisfying* $\sqrt{n}\delta_{crit,\mathcal{G}}^2 \geq \left\|\mathcal{G}\right\|_\infty \mathcal{J}(\delta_{crit,\mathcal{G}}/\left\|\mathcal{G}\right\|_\infty, \mathcal{G})$*, it holds that* $\sqrt{E\left\|\mathcal{H}\mathcal{G}\right\|_{P_n}^2} \lesssim \left\|\mathcal{H}\right\|_\infty \max\left(\left\|\mathcal{G}\right\|_{P_0}, \delta_{crit,\mathcal{G}}\right)$.

Proof. For a function class $\mathcal{F}$, let $\mathcal{J}_n(\delta, \mathcal{F}) := \int_0^\delta \sqrt{\log N(\varepsilon, \mathcal{F}, \|\cdot\|_{P_n})}d\varepsilon$. By Dudley's inequality (Equation 2.1 of Theorem 2.1 in van der Vaart and Wellner [2011]), it holds, for $\delta_n := \left\|\mathcal{G}\mathcal{H}\right\|_{P_n}$, that

$$E\left\|\mathbb{G}_n\right\|_{\mathcal{G}\mathcal{H}} \lesssim E\left[\mathcal{J}_n(\delta_n, \mathcal{G}\mathcal{H})\right].$$

Since $\|g_1h_1\|_{P_n} - \|g_2h_2\|_{P_n} \leq \|h_1 - h_2\|_\infty \|g_1\|_{P_n} + \|g_1 - g_2\|_{P_n}\|\|h_2\|_\infty$, it holds that

$$\log N(\varepsilon, \mathcal{G}\mathcal{H}, \|\cdot\|_{P_n}) \lesssim \log N(\varepsilon/\|\mathcal{H}\|_\infty, \mathcal{G}, \|\cdot\|_{P_n}) + \log N_\infty(\varepsilon/\|\mathcal{G}\|_{P_n}, \mathcal{G}),$$

and, hence,

$$\begin{aligned}
\mathcal{J}_n(\delta, \mathcal{G}\mathcal{H}) &= \int_0^\delta \sqrt{\log N(\varepsilon, \mathcal{G}\mathcal{H}, \|\cdot\|_{P_n})}d\varepsilon \\
&\lesssim \int_0^\delta \sqrt{\log N(\varepsilon/\|\mathcal{H}\|_\infty, \mathcal{G}, \|\cdot\|_{P_n})}d\varepsilon + \int_0^\delta \sqrt{\log N_\infty(\varepsilon/\|\mathcal{G}\|_{P_n}, \mathcal{H})}d\varepsilon \\
&\lesssim \|\mathcal{H}\|_\infty \mathcal{J}(\delta/\|\mathcal{H}\|_\infty, \mathcal{G}) + \|\mathcal{G}\|_{P_n} \mathcal{J}_\infty(\delta/\|\mathcal{G}\|_{P_n}, \mathcal{H}).
\end{aligned}$$

Thus, the identity $E\left\|\mathbb{G}_n\right\|_{\mathcal{G}\mathcal{H}} \lesssim E\left[\mathcal{J}_n(\delta_n, \mathcal{G}\mathcal{H})\right]$ implies that

$$E\left\|\mathbb{G}_n\right\|_{\mathcal{G}\mathcal{H}} \lesssim E\left[\|\mathcal{H}\|_\infty \mathcal{J}(\delta_n/\|\mathcal{H}\|_\infty, \mathcal{G})\right] + E\left[\|\mathcal{G}\|_{P_n} \mathcal{J}_\infty(\delta_n/\|\mathcal{G}\|_{P_n}, \mathcal{H})\right].$$

As in the proof of Theorem 2.1 of van der Vaart and Wellner [2011], concavity of the maps $(x, y) \mapsto \sqrt{y}\mathcal{J}_\infty(\sqrt{x/y}, \mathcal{H})$ and Jenson's inequality implies that

$$E\left\|\mathbb{G}_n\right\|_{\mathcal{G}\mathcal{H}} \lesssim \|\mathcal{H}\|_\infty \mathcal{J}(\sqrt{E\delta_n^2}/\|\mathcal{H}\|_\infty, \mathcal{G}) + \sqrt{E\|\mathcal{G}\|_{P_n}^2} \mathcal{J}_\infty(\sqrt{E\delta_n^2}/\sqrt{E\|\mathcal{G}\|_{P_n}^2}, \mathcal{H}).$$

The first bound then follows from the definition of δ_n. The remaining norm bounds are an immediate consequence of Theorem 2.1 and Theorem 2.2 in van de Geer [2014].

□

The following lemma is an immediate corollary of the above lemma and is used directly in our proofs.

Lemma B.6. *Let* $\mathcal{H}$ *be a uniformly bounded function class satisfying* $\mathcal{J}_\infty(\infty, \mathcal{H}) < \infty$ *and let* $\mathcal{G}$ *be a function class with* $\mathcal{J}(\delta, \mathcal{G}) \lesssim \delta\sqrt{k(n)\log(1/\delta)}$ *where* $\log(1/\left\|\mathcal{G}\right\|) + \log(1/\left\|\mathcal{H}\right\|) = O(1/\log n)$*. Then,*

$$\begin{aligned}
E\left\|\mathbb{G}_n\right\|_{\mathcal{H}\mathcal{G}} \lesssim \left\|\mathcal{G}\right\|_{P_0} \mathcal{J}_\infty\left(\max\left\{\left\|\mathcal{H}\right\|_{P_0}, n^{-1/2}\right\}/\left\|\mathcal{G}\right\|_{P_0}, \mathcal{H}\right) \\
+ \left\|\mathcal{H}\right\|_\infty \sqrt{k(n)\log n} \cdot \max\left\{\left\|\mathcal{G}\right\|_{P_0}, \sqrt{k(n)\log n/n}\right\}.
\end{aligned}$$

B.3 Statement and proofs of technical lemmas

B.3.1 Lemmas bounding key excess risk remainder terms

In this section, we use the supporting lemmas of Appendix B.2.2 to obtain bounds for various remainder terms that appear in our proofs.

The following lemma demonstrates that Lipschitz transformations of the debiased outcome regression estimator $\mu^*_{n,j}$, where $j \in [J]$, falls in function class that is, deterministic conditional on the training set $\mathcal{D}_n \backslash \mathcal{D}^j_n$, and has uniform entropy integral at $\delta > 0$ scaling as $\delta\sqrt{k(n)\log(1/\delta)}$.

Lemma B.7. *Assume Condition C1 and work on the good events E1, E3, and E4. Let $T_{n,j} : \mathbb{R} \to \mathbb{R}$ be a uniformly bounded L-Lipschitz continuous function that is deterministic conditional on the data-fold $\mathcal{D}_n \backslash \mathcal{D}^j_n$. There exists a uniformly bounded random function class $\mathcal{E}_{n,j}$ with the following properties on the good events. (i) $T_{n,j}(\mu_{n,j^*}) - T_{n,j}(\mu_0)$ and $T_{n,j}(\mu_{n,j^*}) - T_{n,j}(\mu_{n,j})$ fall in $\mathcal{E}_{n,j}$ almost surely; (ii) $\mathcal{E}_{n,j}$ is deterministic conditional on $\mathcal{D}_n \backslash \mathcal{D}^j_n$; (iii) $\|\mathcal{E}_{n,j}\| \lesssim LMn^{-\beta/(2\beta+1)} + LM\sqrt{k(n)\log n/n}$; and (iv) $\mathcal{J}(\delta, \mathcal{E}_{n,j}) \lesssim \delta\sqrt{k(n)\log(1/\delta)}$.*

Proof of Lemma B.7. We work on Events E1, E3, and E4. By construction, as shown in Algorithm 4, we have the expression $\mu^*_{n,j} = g^{-1}(g \circ \mu_{n,j} + \widehat{\varphi}^\top_{k(n),j}\beta_n)$, where $\mu^*_{n,j}$ depends randomly only through $\beta_n \in \mathbb{R}^{2k(n)}$, conditional on $\mathcal{D}_n \backslash \mathcal{D}^j_n$. Assuming C1 and on Event E4, the mapping $(a, w) \mapsto \widehat{\varphi}^\top_{k(n),j}(a, w)\beta_n$ belongs to a uniformly bounded subset of a random linear space of dimension $2k(n)$. This space has a VC subgraph dimension of $O(k(n))$, as proven in Lemma 2.6.15 of van der Vaart and Wellner [1996]. Additionally, this function class is deterministic conditional on $\mathcal{D}_n \backslash \mathcal{D}^j_n$.

Considering that $T_{n,j}$ and g are Lipschitz-continuous functions, and $\mu_{n,j}$ and μ_0 are uniformly bounded by C1b, the existence of the above function class and Theorem 2.10.20 of van der Vaart and Wellner [1996] imply that $T_{n,j}(\mu^*_{n,j}) - T_{n,j}(\mu_0)$ and $T_{n,j}(\mu^*_{n,j}) - T_{n,j}(\mu_{n,j})$ fall with probability 1 in a single uniformly bounded function space $\mathcal{W}_{n,j}$ with uniform entropy integral $\mathcal{J}(\delta, \mathcal{W}_{n,j}) \lesssim \delta\sqrt{k(n)\log(1/\delta)}$. Again, this function class is deterministic conditional on $\mathcal{D}_n \backslash \mathcal{D}^j_n$.

Let $\mathcal{E}_{n,j}$ be the subset of $\mathcal{W}_{n,j}$ consisting of elements with $L^2(P_0)$-norm less than $2LMn^{-\beta/(2\beta+1)} + 2LM\sqrt{k(n)\log n/n}$, where L is the Lipschitz constant of $T_{n,j}$. In view of Event E3, using the Lipschitz-continuity of $T_{n,j}$, we can deduce that the event $T_{n,j}(\mu^*_{n,j}) - T_{n,j}(\mu_0) \in \mathcal{E}_{n,j}$ occurs almost surely on the good events. Moreover, on the good events, we also have $T_{n,j}(\mu^*_{n,j}) - T_{n,j}(\mu_{n,j}) \in \mathcal{E}_{n,j}$, noting that

$$\begin{aligned}\|T_{n,j}(\mu_{n,j}) - T_{n,j}(\mu^*_{n,j})\| &\leq \|T_{n,j}(\mu_{n,j}) - T_{n,j}(\mu_0)\| + \|T_{n,j}(\mu_0) - T_{n,j}(\mu^*_{n,j})\| \\ &\leq 2LMn^{-\beta/(2\beta+1)} + 2LM\sqrt{k(n)\log n/n}.\end{aligned}$$

□

The next few lemmas bound the localized suprema of various empirical process remainders that appear in our proofs. For this, we recall the sup-norm coupling exponent $\alpha > 0$ of C5 and the notation $r^*_n := n^{-\beta/(2\beta+1)} + \sqrt{k(n)\log n/n}$ and $s^*_n := n^{-\gamma/(2\gamma+1)} + r^*_n$. We also recall that $\mathbb{I}_{A_n}$ is the indicator that events E1- E4 occur.

Lemma B.8. *Let $\mathcal{G}$ be any uniformly bounded function class with $\mathcal{J}_\infty(\infty, \mathcal{G}) < \infty$ and $m \in \{1, 2\}$. Under the conditions of Theorem 3.4, it holds, for all $\delta \geq \|\mathcal{G}\| + n^{-1/2}$ and $\delta_\infty \geq \|\mathcal{G}\|_\infty$, that*

$$E_0^n \left[\mathbb{I}_{A_n} \sup_{\phi \in \mathcal{G}} \left| (\overline{P}_n - \overline{P}_0) \left\{ \overline{\Delta}^{(m)}_{\pi_{n,\diamond}\mu^*_{n,\diamond}}(\cdot\,, \phi) - \overline{\Delta}^{*(m)}_{\pi_{n,\diamond}\mu_{n,\diamond}}(\cdot\,, \phi) \right\} \right| \right]$$
$$\lesssim \frac{r_n^* \mathcal{J}_\infty\left(\delta / r_n^*, \mathcal{G}\right)}{\sqrt{n}} + \delta_\infty r_n^* \sqrt{k(n) \log n / n}.$$

Proof of Lemma B.8. By our notation, we have, for $\phi \in \mathcal{G}$, that

$$(\overline{P}_n - \overline{P}_0) \left\{ \overline{\Delta}^{(m)}_{\pi_{n,\diamond}\mu^*_{n,\diamond}}(\cdot\,, \phi) - \overline{\Delta}^{*(m)}_{\pi_{n,\diamond}\mu_{n,\diamond}}(\cdot\,, \phi) \right\}$$
$$= \frac{1}{J} \sum_{j \in [J]} (P_{n,j} - P_0) \left\{ \overline{\Delta}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(\cdot\,, \phi) - \overline{\Delta}^{*(m)}_{\pi_{n,j},\mu_{n,j}}(\cdot\,, \phi) \right\}. \tag{B.1}$$

By the triangle inequality for a given $j \in [J]$, it suffices to bound

$$E_0^n \left[\mathbb{I}_{A_n} \sup_{\phi \in \mathcal{G}} \left| (P_{n,j} - P_0) \left\{ \overline{\Delta}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(\cdot\,, \phi) - \overline{\Delta}^{*(m)}_{\pi_{n,j},\mu_{n,j}}(\cdot\,, \phi) \right\} \right| \right].$$

To do so, we apply, conditionally on $\mathcal{D}_n \backslash \mathcal{D}_n^j$, the maximal inequality of Lemma B.6, where we use Lemma B.7 to control the randomness of $\mu^*_{n,j}$. To this end, for $\phi \in \mathcal{G}$, $j \in [J]$, and $o \in \mathcal{O}$, we have, by definition, that

$$\overline{\Delta}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(o\,, \phi) - \overline{\Delta}^{*(m)}_{\pi_{n,j},\mu_{n,j}}(o\,, \phi) = \frac{y - \mu^*_{n,j}(a, w)}{\pi_{n,j}(a \,|\, w)} \left\{ H_{m,\mu^*_{n,j}}(a, w) - H_{m,\mu_{n,j}}(a, w) \right\} \{\phi(w)\},$$

where we recall that $H_{m,\mu}(a, w) := c_{a,m}(\dot{g}_m \circ \mu)(a, w)$. By Lipschitz continuity of the functions $\dot{g}_1$ and $\dot{g}_2$ in the definition of $H_{m,\mu}$, we have, for some constant $L > 0$, each $o \in \mathcal{O}$, and any map $(a, w) \mapsto \mu(a, w)$, that

$$\left| H_{m,\mu}(a, w) - H_{m,\mu_{n,j}}(a, w) \right| \lesssim L |\mu(a, w) - \mu_{n,j}(a, w)|.$$

As a consequence, defining the Lipschitz continuous transformation $\mu \mapsto T_{n,j}(\mu) := H_{m,\mu}$, we can write $H_{m,\mu} - H_{m,\mu_{n,j}} = T_{n,j}(\mu) - T_{n,j}(\mu_{n,j})$ where $T_{n,j}$ is deterministic given $\mathcal{D}_n \backslash \mathcal{D}_n^j$. Hence, by Lemma B.7, on the event A_n, there exists a function class $\mathcal{E}_{n,j}$ that is deterministic conditional on $\mathcal{D}_n \backslash \mathcal{D}_n^j$ with $H_{m,\mu^*_{n,j}} - H_{m,\mu_{n,j}} \in \mathcal{E}_{n,j}$ almost surely. Moreover, on the event A_n, this function class satisfies $\mathcal{J}(\delta, \mathcal{E}_{n,j}) \lesssim \delta \sqrt{k(n) \log(1/\delta)}$ and $\|\mathcal{E}_{n,j}\| \lesssim r_n^*$.

In the following, we work on the event A_n. Define the function class,

$$\mathcal{W}_{n,j} := \left\{ o \mapsto (y - \mu(a, w)) / \pi_{n,j}(a \,|\, w) \cdot \left\{ H_{m,\mu}(a, w) - H_{m,\mu_{n,j}}(a, w) \right\} : \mu \in \mathcal{E}_{n,j} \right\},$$

which is deterministic conditional on $\mathcal{D}_n \backslash \mathcal{D}_n^j$. Note, by C1a and boundedness of $\mathcal{E}_{n,j}$, that $\|\mathcal{W}_{n,j}\| \lesssim \|\mathcal{E}_{n,j}\| \lesssim r_n^*$. Moreover, for each $\phi \in \mathcal{G}$, the map $o \mapsto \Delta^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(o\,, \phi) - \overline{\Delta}^{*(m)}_{\pi_{n,j},\mu_{n,j}}(o\,, \phi)$ falls almost surely in the function class $\mathcal{G}\mathcal{W}_{n,j} := \{\phi g : \phi \in \mathcal{G}, g \in \mathcal{W}_{n,j}\}$.

Conditions C1a and C1b along with Theorem 2.10.20. of van der Vaart and Wellner [1996] imply, conditional on $\mathcal{D}_n \backslash \mathcal{D}_n^j$, that $\mathcal{J}(\delta, \mathcal{W}_{n,j}) \lesssim \mathcal{J}(\delta, \mathcal{E}_{n,j}) \lesssim \delta\sqrt{k(n)\log n}$ for any $\delta > 1/\sqrt{n}$.

We are now in a position to bound the empirical process term of the right-hand side of (B.1). By the above construction, we have that

$$E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{G}}\left|(P_{n,j}-P_0)\left\{\overline{\Delta}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(\cdot\,,\phi)-\overline{\Delta}^{*(m)}_{\pi_{n,j},\mu_{n,j}}(\cdot\,,\phi)\right\}\right| \mid \mathcal{T}_n^j\right]$$
$$\leq E_0^n\left[\mathbb{I}_{A_n}\sup_{hg\in\mathcal{G}\mathcal{W}_{n,j}}|(P_{n,j}-P_0)\,hg| \mid \mathcal{T}_n^j\right]$$

Applying Lemma B.6 conditional on $\mathcal{D}_n \backslash \mathcal{D}_n^j$ and on the event A_n with $\mathcal{G} := \mathcal{G}$ and $\mathcal{W} := \mathcal{W}_{n,j}$, we find

$$E_0^n\left[\mathbb{I}_{A_n}\sup_{hg\in\mathcal{G}\mathcal{W}_{n,j}}|(P_{n,j}-P_0)\,hg| \mid \mathcal{T}_n^j\right]$$
$$\lesssim E_0^n\left[\mathbb{I}_{A_n}\left\{n^{-1/2}\left\|\mathcal{W}_{n,j}\right\|\mathcal{J}_\infty\left(\max\left\{\|\mathcal{G}\|,n^{-1/2}\right\}/\left\|\mathcal{W}_{n,j}\right\|,\mathcal{G}\right) \mid \mathcal{T}_n^j\right\}\right]$$
$$+E_0^n\left[\mathbb{I}_{A_n}\left\{\|\mathcal{G}\|_\infty\sqrt{k(n)\log n/n}\cdot\max\left\{\left\|\mathcal{W}_{n,j}\right\|_{P_0},\sqrt{k(n)\log n/n}\right\}\right\} \mid \mathcal{T}_n^j\right].$$

Next, using that $\|\mathcal{W}_{n,j}\| \lesssim r_n^*$ on the event A_n and taking expectations of both sides of the previous inequality, we find

$$E_0^n\left[\mathbb{I}_{A_n}\sup_{hg\in\mathcal{G}\mathcal{W}_{n,j}}|(P_{n,j}-P_0)\,hg|\right] \lesssim n^{-1/2}r_n^*\mathcal{J}_\infty\left(\max\{\|\mathcal{G}\|,n^{-1/2}\}/r_n^*,\mathcal{G}\right)$$
$$+\,r_n^*\|\mathcal{G}\|_\infty\sqrt{k(n)\log n/n}.$$

Summing the above over $j \in [J]$ and applying the reverse triangle inequality, we obtain the desired result.

□

Lemma B.9. *Let $\mathcal{G}$ be any uniformly bounded function class with $\mathcal{J}_\infty(\infty,\mathcal{G}) < \infty$ and $m \in \{1,2\}$. Assume the conditions of Theorem 3.4 and Events E1- E4. For all $\delta \geq \|\mathcal{G}\| + n^{-1/2}$ and $\delta_\infty \geq \|\mathcal{G}\|_\infty$, it holds that*

$$E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{G}}\left|(\overline{P}_n-\overline{P}_0)\left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\phi)-\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu_0}(\cdot\,;\phi)\right\}\right|\right]$$
$$\lesssim \frac{r_n^*\mathcal{J}_\infty\left(\delta/r_n^*,\mathcal{G}\right)}{\sqrt{n}}+\delta_\infty r_n^*\sqrt{k(n)\log n/n}.$$

Proof of Lemma B.9. This proof follows from minor modifications to the proof of Lemma B.8. As before, by the triangle inequality, it suffices to bound, for $j \in [J]$, the term

$$E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{G}}\left|(P_{n,j}-P_0)\left\{\overline{\Delta}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(\cdot\,;\phi)-\overline{\Delta}^{(m)}_{\pi_{n,j},\mu_0}(\cdot\,;\phi)\right\}\right|\right].$$

Arguing as in the proof of Lemma B.8 and applying C1a and C1b, we can show, for each

$o \in \mathcal{O}$, that

$$\left|\overline{\Delta}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(o\,;\phi) - \overline{\Delta}^{(m)}_{\pi_{n,j},\mu_0}(o\,,\phi)\right| \lesssim \left|\{\mu^*_{n,j}(a,w) - \mu_0(a,w)\}\phi(w)\right|.$$

On Event E3 and a similar argument as in the proof of Lemma B.8, we can find a product function class $\mathcal{GW}_{n,j} = \{hw : \phi \in \mathcal{G}, w \in \mathcal{W}_{n,j}\}$ that contains $\overline{\Delta}_{\pi_{n,j},\mu^*_{n,j}}(\cdot\,;\phi) - \overline{\Delta}_{\pi_{n,j},\mu_0}(\cdot\,;\phi)$ with probability 1. Moreover, on event A_n and by Lemma B.7, the function class $\mathcal{W}_{n,j}$ can be chosen deterministic conditional on $\mathcal{D}_n \backslash \mathcal{D}^j_n$ and to satisfy $\|\mathcal{W}_{n,j}\| \lesssim r^*_n$ and $\mathcal{J}(\delta, \mathcal{W}_{n,j}) \lesssim \delta\sqrt{k(n)\log(1/\delta)} \lesssim \delta\sqrt{k(n)\log n}$ for any $\delta > n^{-1/2}$. Arguing as in the proof of Lemma B.8, we find

$$\begin{aligned} E^n_0\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{G}}\left|(P_{n,j}-P_0)\left\{\overline{\Delta}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(\cdot\,;\phi) - \overline{\Delta}^{(m)}_{\pi_{n,j},\mu_0}(\cdot\,;\phi)\right\}\right|\right] &\leq E^n_0\left[\mathbb{I}_{A_n}\sup_{hg\in\mathcal{GW}_{n,j}}|(P_{n,j}-P_0)\,hg|\right] \\ &\lesssim \frac{r^*_n\mathcal{J}_\infty\left(\delta/r^*_n,\mathcal{G}\right)}{\sqrt{n}} + \delta_\infty r^*_n\sqrt{k(n)\log n/n} \end{aligned}$$

The result follows, noting that the final bound of the above display does not depend on j. □

Lemma B.10. *Let $\mathcal{G}$ be any uniformly bounded function class with $\mathcal{J}_\infty(\infty,\mathcal{G}) < \infty$ and $m \in \{1,2\}$. Assume the conditions of Theorem 3.4 and Events E1- E4. For any $\delta \geq \|\mathcal{G}\| + n^{-1/2}$, it holds that*

$$E^n_0\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{G}}\left|(\overline{P}_n - \overline{P}_0)\left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu_0}(\cdot\,;\phi)\right\}\right|\right] \lesssim n^{-1/2}\mathcal{J}_\infty\left(\delta,\mathcal{G}\right)$$

Proof of Lemma B.10. As in the proof of Lemma B.9, it suffices to bound

$$E^n_0\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{G}}\left|(P_{n,j}-P_0)\left\{\overline{\Delta}^{(m)}_{\pi_{n,j},\mu_0}(\cdot\,;\phi)\right\}\right|\right].$$

From the proof of Lemma B.9, we know that $\phi(\,\cdot\,) \mapsto \left\{\overline{\Delta}^{(m)}_{\pi_{n,j},\mu_0}(\cdot\,;\phi)\right\}$ is a Lipschitz transformation of $\phi(\,\cdot\,)$. Moreover, this Lipschitz map is fixed conditional on the training set $\mathcal{D}_n \backslash \mathcal{D}^j_n$. Also, on event A_n, the Lipschitz map and its constant are uniformly bounded since μ_0 and $\pi^{-1}_{n,j}$ are uniformly bounded by C1b and C1a. Thus, on A_n, we have that $\left\{\overline{\Delta}^{(m)}_{\pi_{n,j},\mu_0}(\cdot\,;\phi) : \phi \in \mathcal{H}\right\}$ falls in a random bounded function class that is deterministic conditional on $\mathcal{D}_n \backslash \mathcal{D}^j_n$ and, by Lemma B.2, has sup-norm entropy integral bounded by $C\mathcal{J}_\infty(\delta,\mathcal{G})$, up to a constant. The result then follows from a direct application of Lemma B.4, applied on event A_n and conditionally on $\mathcal{D}_n \backslash \mathcal{D}^j_n$ as in Lemma B.8.

□

In the following lemmas, we denote as shorthand $\left[F - \widetilde{F}\right][\theta_1 - \theta_2] := \left[F(\theta_1) - \widetilde{F}(\theta_1)\right] - \left[F(\theta_2) - \widetilde{F}(\theta_2)\right]$ for any two functionals $F, \widetilde{F} \in \ell^\infty(\mathcal{F})$ and $\theta_1, \theta_2 \in \mathcal{F}$. We recall that $s^*_n := n^{-\gamma/(2\gamma+1)} + n^{-\beta/(2\beta+1)} + \sqrt{k(n)\log n/n}$.

Lemma B.11. *Assume the conditions of Theorem 3.4 and Events E1- E4. For $\delta > n^{-1/2}$, we have*

$$E_0^n\Big[\mathbb{I}_{A_n}\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta}\Big|\Big[(\overline{P}_n-\overline{P}_0)\left(L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}-L_{\pi_0,\mu_0}\right)\Big][\theta_1-\theta_2]\Big|\Big]$$
$$\lesssim n^{-1/2}s_n^*\mathcal{J}_\infty\left(\delta/s_n^*,\mathcal{F}\right)+r_n^*\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta}\|\theta\|_\infty\sqrt{k(n)\log n/n}.$$

Similarly, we have

$$E_0^n\Big[\mathbb{I}_{A_n}\sup_{\theta\in\mathcal{F}:\|\theta\|\le\delta}\Big|(\overline{P}_n-\overline{P}_0)\left\{L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\theta)-L_{\pi_0,\mu_0}(\cdot\,,\theta)\right\}\Big|\Big]$$
$$\lesssim n^{-1/2}s_n^*\mathcal{J}_\infty\left(\delta/s_n^*,\mathcal{F}\right)+r_n^*\sup_{\theta\in\mathcal{F}:\|\theta\|\le\delta}\|\theta\|_\infty\sqrt{k(n)\log n/n}.$$

Proof of Lemma B.11. For $\phi\in\mathcal{H}$ and $m\in\{1,2\}$, we denote

$$\overline{L}^{(m)}_\mu(\cdot,\phi):=\phi(\cdot)\cdot\sum_{a\in\mathcal{A}}c_{a,m}(g_m\circ\mu)(a,\cdot)$$
$$\overline{L}^{(m)}_{\pi,\mu}(\cdot,\phi):=\overline{L}^{(m)}_\mu(\cdot,\phi)+\overline{\Delta}^{(m)}_{\pi,\mu}(\cdot,\phi).$$

Observe that the function h_m in the definition of the loss $L_{\pi,\mu}$ and $\mathcal{H}$ are L-Lipschitz continuous for some $L>0$. Hence, by the triangle inequality and linearity of $L_{\pi,\mu}(\cdot,\theta)$ viewed as a mapping in $h_1\circ\theta$ and $h_2\circ\theta$, we find

$$E_0^n\left[\mathbb{I}_{A_n}\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta}\Big|\Big[(\overline{P}_n-\overline{P}_0)\left(L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}-L_{\pi_0,\mu_0}\right)\Big][\theta_1-\theta_2]\Big|\right]$$
$$\lesssim\max_{j\in[J],m\in\{1,2\}}E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi_1,\phi_2\in\mathcal{H}:\|\phi_1-\phi_2\|\le L\delta}\Big|\Big[(P_{n,j}-P_0)\left(\overline{L}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}-\overline{L}^{(m)}_{\pi_0,\mu_0}\right)(\cdot,\phi_1-\phi_2)\Big]\Big|\right].$$

To bound the right-hand side, it suffices to bound the following for a given $j\in[J]$ and $m\in\{1,2\}$:

$$E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi_1,\phi_2\in\mathcal{H}:\|\phi_1-\phi_2\|\le L\delta}\left|(P_{n,j}-P_0)\left(\overline{L}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}-\overline{L}^{(m)}_{\pi_0,\mu_0}\right)(\cdot,\phi_1-\phi_2)\right|\right].$$

Towards this goal, we have the expansion:

$$(P_{n,j}-P_0)\left\{\overline{L}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(\cdot\,,\phi_1-\phi_2)-\overline{L}^{(m)}_{\pi_0,\mu_0}(\cdot\,,\phi_1-\phi_2)\right\}$$
$$=(P_{n,j}-P_0)\left\{\overline{L}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(\cdot\,,\phi_1-\phi_2)-\overline{L}^{(m)}_{\pi_{n,j},\mu_0}(\cdot\,,\phi_1-\phi_2)\right\}$$
$$+(P_{n,j}-P_0)\left\{\overline{L}^{(m)}_{\pi_{n,j},\mu_0}(\cdot\,,\phi_1-\phi_2)-\overline{L}^{(m)}_{\pi_0,\mu_0}(\cdot\,,\phi_1-\phi_2)\right\}$$
$$\le\text{(I)}+\text{(II)}+\text{(III)},$$

where, noting that $\overline{L}^{(m)}_{\pi,\mu}(\cdot,\phi)=\overline{L}^{(m)}_\mu(\cdot,\phi)+\overline{\Delta}^{(m)}_{\pi,\mu}(\cdot,\phi)$, we define:

$$\text{(I)}:=\sup_{\phi_1,\phi_2\in\mathcal{H}:\|\phi_1-\phi_2\|\le L\delta}\left|(P_{n,j}-P_0)\left\{\overline{L}^{(m)}_{\mu^*_{n,j}}(\cdot\,,\phi_1-\phi_2)-\overline{L}^{(m)}_{\mu_0}(\cdot\,,\phi_1-\phi_2)\right\}\right|;$$
$$\text{(II)}:=\sup_{\phi_1,\phi_2\in\mathcal{H}:\|\phi_1-\phi_2\|\le L\delta}\left|(P_{n,j}-P_0)\left\{\overline{\Delta}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(\phi_1-\phi_2)-\overline{\Delta}^{(m)}_{\pi_{n,j},\mu_0}(\phi_1-\phi_2)\right\}\right|;$$

$$\text{(III)} := \sup_{\phi_1,\phi_2\in\mathcal{H}:\|\phi_1-\phi_2\|\le L\delta} \left|(P_{n,j}-P_0)\left\{\overline{L}^{(m)}_{\pi_{n,j},\mu_0}(\cdot\,,\phi_1-\phi_2)-\overline{L}^{(m)}_{\pi_0,\mu_0}(\cdot\,,\phi_1-\phi_2)\right\}\right|.$$

We proceed by bounding in expectation each of the above terms on event A_n. To bound (I), take $\phi_1,\phi_2\in\mathcal{H}$ with $\|\phi_1-\phi_2\|\le L\delta$ and observe that

$$\overline{L}^{(m)}_{\mu^*_{n,j}}(\cdot\,,\phi_1-\phi_2)-\overline{L}^{(m)}_{\mu_0}(\cdot\,,\phi_1-\phi_2)=(\phi_1-\phi_2)\sum_{a\in\mathcal{A}}c_{a,m}\left\{g_m\circ\mu^*_{n,j}(a,\cdot)-g_m\circ\mu_0(a,\cdot)\right\}.$$

Arguing as in the proof of Lemma B.8, using C1, and applying Lemma B.7, on event A_n, we can almost surely find a function class $\mathcal{E}_{n,j}$ containing $\{g_m\circ\mu^*_{n,j}(a,\cdot)-g_m\circ\mu_0(a,\cdot):a\in\mathcal{A}\}$ that is deterministic conditional on $\mathcal{D}_n\backslash\mathcal{D}^j_n$. Moreover, the class $\mathcal{E}_{n,j}$ can be chosen such that $\mathcal{J}(\delta,\mathcal{E}_{n,j})\lesssim\delta\sqrt{\log(1/\delta)k(n)}\lesssim\delta\sqrt{k(n)\log n}$ and $\|\mathcal{E}_{n,j}\|\lesssim r^*_n$. Hence, on the event A_n, $\overline{L}^{(m)}_{\mu^*_{n,j}}(\cdot\,,\phi_1-\phi_2)-\overline{L}^{(m)}_{\mu_0}(\cdot\,,\phi_1-\phi_2)$ is contained in the product class $\mathcal{GE}_{n,j}:=\{gh:g\in\mathcal{G},h\in\mathcal{E}_{n,j}\}$ where $\mathcal{G}:=\{\phi_1-\phi_2:\phi_1,\phi_2\in\mathcal{H},\|\phi_1-\phi_2\|\le L\delta\}$. Arguing as in Lemma B.9 and working on event A_n, we can apply Theorem 2.10.20 in van der Vaart and Wellner [1996] and Lemma B.6 to bound (I) in expectation as

$$E^n_0\left[\mathbb{I}_{A_n}\cdot\text{(I)}\right]\lesssim n^{-1/2}r^*_n\mathcal{J}_\infty\left(\delta/r^*_n,\mathcal{F}\right)+r^*_n\delta_\infty\sqrt{k(n)\log n/n}.$$

Next, to bound (II), we have, as a direct consequence of Lemma B.9 applied to the function class $\{\phi_1,\phi_2\in\mathcal{H}:\|\phi_1-\phi_2\|\le L\delta\}$, that

$$E^n_0\left[\mathbb{I}_{A_n}\cdot\text{(II)}\right]\lesssim n^{-1/2}r^*_n\mathcal{J}_\infty\left(\delta/r^*_n,\mathcal{F}\right)+r^*_n\delta_\infty\sqrt{k(n)\log n/n}.$$

Lastly, to bound (III), we use that

$$\begin{aligned}&\sup_{\phi_1,\phi_2\in\mathcal{H}:\|\phi_1-\phi_2\|\le L\delta}\left|(P_{n,j}-P_0)\left\{\overline{L}^{(m)}_{\pi_{n,j},\mu_0}(\cdot\,,\phi_1-\phi_2)-\overline{L}^{(m)}_{\pi_0,\mu_0}(\cdot\,,\phi_1-\phi_2)\right\}\right|\\&\quad=\sup_{\phi_1,\phi_2\in\mathcal{H}:\|\phi_1-\phi_2\|\le L\delta}\left|(P_{n,j}-P_0)\left\{\overline{\Delta}^{(m)}_{\pi_{n,j},\mu_0}(\cdot\,,\phi_1-\phi_2)-\overline{\Delta}^{(m)}_{\pi_0,\mu_0}(\cdot\,,\phi_1-\phi_2)\right\}\right|.\end{aligned}$$

We will argue along similar lines as the proof of Lemma B.9. We have $\overline{\Delta}^{(m)}_{\pi_{n,j},\mu_0}(\cdot\,,\phi_1-\phi_2)-\overline{\Delta}^{(m)}_{\pi_0,\mu_0}(\cdot\,,\phi_1-\phi_2)$ is fixed conditional on $\mathcal{D}_n\backslash\mathcal{D}^j_n$. Moreover, by C1a and C1b, this function can be expressed as a Lipschitz transformation of the function class $\mathcal{W}_{n,j}:=\{(\pi_{n,j}-\pi_0)(a\,|\,\cdot)(\phi_1-\phi_2):a\in\mathcal{A}:\phi_1,\phi_2\in\mathcal{H}\}$. Denote the rate of C2a for $\|\pi_{n,j}-\pi_0\|$ as $\varepsilon_{n,j}:=n^{-\gamma/(2\gamma+1)}$. Working on Event E2 and applying Lemma B.4 conditionally on $\mathcal{D}_n\backslash\mathcal{D}^j_n$ to $\mathcal{W}_{n,j}$, we obtain the bound:

$$\begin{aligned}E^n_0\left[\mathbb{I}_{A_n}\cdot\text{(III)}\right]&\lesssim n^{-1/2}\int_0^{\varepsilon_{n,j}}\sqrt{\log N_\infty(\varepsilon,\mathcal{W}_{n,j})}d\varepsilon\\&\lesssim n^{-1/2}\varepsilon_{n,j}\mathcal{J}_\infty\left(\delta/\varepsilon_{n,j},\mathcal{F}\right)\\&\lesssim n^{-1/2}s^*_n\mathcal{J}_\infty\left(\delta/s^*_n,\mathcal{F}\right),\end{aligned}$$

where the second inequality follows from $N_\infty(\varepsilon,\mathcal{W}_{n,j})\lesssim N_\infty(\varepsilon/\varepsilon_{n,j},\mathcal{F})$ and a change of variables – see the proof of Lemma B.6 for a related argument.

Now, we note that $n^{-1/2}r^*_n\mathcal{J}_\infty\left(\delta/r^*_n,\mathcal{F}\right)\lesssim n^{-1/2}s^*_n\mathcal{J}_\infty\left(\delta/s^*_n,\mathcal{F}\right)$ since $r^*_n\le s^*_n$. Using this, combining the above bounds for (I), (II), and (III), and applying the triangle inequality, we finally obtain:

$$E^n_0\left[\mathbb{I}_{A_n}\cdot\text{(I)}\right]+E^n_0\left[\mathbb{I}_{A_n}\cdot\text{(II)}\right]+E^n_0\left[\mathbb{I}_{A_n}\cdot\text{(III)}\right]\lesssim n^{-1/2}s^*_n\mathcal{J}_\infty\left(\delta/s^*_n,\mathcal{F}\right)+r^*_n\delta_\infty\sqrt{k(n)\log n/n},$$

as desired.

□

The next lemmas bound several second-order remainders that appear in our proofs. Recall that $\mathbb{I}_{g_1,g_2}$ is the indicator that takes the value 0 if both g_1 and g_0 are the identity function.

Lemma B.12. *Assume the conditions of Theorem 3.4. On event A_n, for any $\theta_1, \theta_2 \in \mathcal{F}$ and $\delta > n^{-1/2}$, we have*

$$\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta} \left| R_{\pi_{n,\diamond}\mu^*_{n,\diamond}}(\theta_2) - R_{\pi_0,\mu_0}(\theta_2) - R_{\pi_{n,\diamond}\mu^*_{n,\diamond}}(\theta_1) + R_{\pi_0,\mu_0}(\theta_1) \right| \lesssim \sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta} \left\{ r_n^* n^{-\gamma/(2\gamma+1)} + \mathbb{I}_{g_1,g_2} \cdot (r_n^*)^2 \right\}.$$

Similarly, on event A_n, we have

$$\sup_{\theta\in\mathcal{F}:\|\theta\|\le\delta} \left| R_{\pi_{n,\diamond}\mu^*_{n,\diamond}}(\theta) - R_{\pi_0,\mu_0}(\theta) \right| \lesssim \sup_{\theta\in\mathcal{F}:\|\theta\|\le\delta} \|\theta\|_\infty \left\{ r_n^* n^{-\gamma/(2\gamma+1)} + \mathbb{I}_{g_1,g_2} \cdot (r_n^*)^2 \right\}.$$

Proof of Lemma B.12. We work on the event A_n, on which, by definition, events E1-E4 also occur. As in the proof of Lemma B.11, for $\phi \in \mathcal{H}$ and $m \in \{1,2\}$, we denote $\overline{L}^{(m)}_{\pi,\mu}(\cdot,\phi) := \phi \cdot \sum_{a\in\mathcal{A}} c_{a,m}(g_m \circ \mu)(a,\cdot) + \overline{\Delta}^{(m)}_{\pi,\mu}(\cdot,\phi)$. To begin, observe that

$$\begin{aligned} R_{\pi_{n,\diamond}\mu^*_{n,\diamond}}(\theta) - R_{\pi_0,\mu_0}(\theta) &= J^{-1} \sum_{j\in[J]} P_0 \left\{ L_{\pi_{n,j},\mu^*_{n,j}}(\cdot,\theta) - L_{\pi_0,\mu_0}(\cdot,\theta) \right\} \\ &= J^{-1} \sum_{j\in[J]} \sum_{m\in\{1,2\}} P_0 \left\{ \overline{L}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(\cdot, h_m\circ\theta) - \overline{L}^{(m)}_{\mu_0}(\cdot, h_m\circ\theta) \right\}, \end{aligned}$$

where, for each $m \in \{1,2\}$, the right-hand side is linear in $h_m \circ \theta$. Moreover, since each h_m is Lipschitz continuous, we have $\|\theta_1 - \theta_2\| \le \delta$ for $\theta_1, \theta_2 \in \mathcal{F}$ implies, for some $L > 0$, that $\|h_m \circ \theta_1 - h_m \circ \theta_2\| \le L\delta$. Hence, taking the supremum over the previous display, we find that

$$\begin{aligned} &\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta} \left| R_{\pi_{n,\diamond}\mu^*_{n,\diamond}} - R_{\pi_0,\mu_0} - R_{\pi_{n,\diamond}\mu^*_{n,\diamond}} + R_{\pi_0,\mu_0} \right| [\theta_1 - \theta_2] \\ &\lesssim \max_{j\in[J],m\in\{1,2\}} \sup_{\phi_1,\phi_2\in\mathcal{H}:\|\phi_1-\phi_2\|\le L\delta} \left| P_0 \left\{ \overline{L}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(\cdot,\phi_1-\phi_2) - \overline{L}^{(m)}_{\mu_0}(\cdot,\phi_1-\phi_2) \right\} \right|, \end{aligned}$$

It suffices to bound, for a given $j \in [J]$ and $m \in \{1,2\}$,

$$P_0 \left\{ \overline{L}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(\cdot,\phi_1-\phi_2) - \overline{L}^{(m)}_{\mu_0}(\cdot,\phi_1-\phi_2) \right\}$$

where $\phi_1, \phi_2 \in \mathcal{H}$ are such that $\|\phi_1 - \phi_2\| \le L\delta$.

Plugging in the definition of the loss functions, we find for $\phi := \phi_1 - \phi_2$ that

$$\begin{aligned} &P_0 \left\{ \overline{L}^{(m)}_{\pi_{n,j},\mu^*_{n,j}}(\cdot,\phi) - \overline{L}^{(m)}_{\mu_0}(\cdot,\phi) \right\} \\ &= E_0 \left[\overline{L}^{(m)}_{\mu^*_{n,j}}(O,\,\phi) - \overline{L}^{(m)}_{\mu_0}(O\,,\phi) \,|\, \mathcal{D}_n \right] \\ &\quad + \sum_{a\in\mathcal{A}} E_0 \left[\frac{1(A=a)}{\pi_{n,j}(A\,|\,W)} \phi(W) c_{a,m}(\dot{g}_m \circ \mu^*_{n,j})(a,W) \left\{ Y - \mu^*_{n,j}(A,W) \right\} \,|\, \mathcal{D}_n \right] \end{aligned}$$

$$
\begin{aligned}
&= E_0\left[\overline{L}^{(m)}_{\mu^*_{n,j}}(O\,,\phi) - \overline{L}^{(m)}_{\mu_0}(O\,,\phi)\,|\,\mathcal{D}_n\right] \\
&\quad + \sum_{a\in\mathcal{A}} E_0\left[\frac{\pi_0(a\,|\,W)}{\pi_{n,j}(a\,|\,W)}\phi(W)c_{a,m}(\dot{g}_m\circ\mu^*_{n,j})(a,W)\left\{\mu_0(a,W)-\mu^*_{n,j}(a,W)\right\}\,|\,\mathcal{D}_n\right] \\
&= \text{(I)} + \text{(II)},
\end{aligned}
$$

where we denote:

$$
\begin{aligned}
\text{(I)} &:= E_0\left[\overline{L}^{(m)}_{\mu^*_{n,j}}(O,\phi) - \overline{L}^{(m)}_{\mu_0}(O\,,\phi)\,|\,\mathcal{D}_n\right] \\
&\quad + \sum_{a\in\mathcal{A}} E_0\left[\phi(W)c_{a,m}(\dot{g}_m\circ\mu^*_{n,j})(a,W)\left\{\mu_0(a,W)-\mu^*_{n,j}(a,W)\right\}\,|\,\mathcal{D}_n\right], \\
\text{(II)} &:= \sum_{a\in\mathcal{A}} E_0\left[\phi(W)c_{a,m}(\dot{g}_m\circ\mu^*_{n,j})(a,W)\left\{1-\frac{\pi_0(a\,|\,W)}{\pi_{n,j}(a\,|\,W)}\right\}\left\{\mu_0(a,W)-\mu^*_{n,j}(a,W)\right\}\,|\,\mathcal{D}_n\right].
\end{aligned}
$$

Applying Cauchy-Schwarz, C1b, C1a, and C4, term (II) can be upper bounded in absolute value by

$$
\begin{aligned}
|\text{(II)}| &\lesssim \max_{a\in\mathcal{A}}\max_{j\in[J]} P_0\left|\phi\{\mu_0(a,\cdot)-\mu_{n,j}(a,\cdot)\}\{\pi_0(a\,|\,\cdot)-\pi_{n,j}(a\,|\,\cdot)\}\right| \\
&\lesssim \max_{j\in[J]} P_0\left|\phi\{\mu_0-\mu_{n,j}\}\{\pi_0-\pi_{n,j}\}\right|,
\end{aligned}
$$

where the final inequality follows from Condition C1a and finiteness of $\mathcal{A}$. To bound term (I), note

$$
\begin{aligned}
\text{(I)} \;&= \sum_{a\in\mathcal{A}} E_0\left[\phi(W)c_{a,m}\left\{(g_m\circ\mu^*_{n,j})(a,W)-(g_m\circ\mu_0)(a,W)\right\}\,|\,\mathcal{D}_n\right] \\
&\quad - \sum_{a\in\mathcal{A}} E_0\left[\phi(W)c_{a,m}(\dot{g}_m\circ\mu^*_{n,j})(a,W)\left\{\mu^*_{n,j}(a,W)-\mu_0(a,W)\right\}\,|\,\mathcal{D}_n\right].
\end{aligned}
$$

First, note that the right-hand side simplifies to zero in the case where $\mathbb{I}_{g_1,g_2}=0$, since $g_m\circ\mu=\mu$ and $\dot{g}_m\circ\mu=1$. Recall that g_m is twice differentiable with Lipschitz continuous derivative. Hence, when $\mathbb{I}_{g_1,g_2}=1$, using C4, C1b, C1a, and a bound on the remainder of a first-order Taylor expansion of g_m, we can continue the previous display as:

$$
\begin{aligned}
\text{(I)} &= \sum_{a\in\mathcal{A}} E_0\Bigg[\phi(W)c_{a,m}\Big\{(g_m\circ\mu^*_{n,j})(a,W)-(g_m\circ\mu_0)(a,W) \\
&\qquad\qquad - (\dot{g}_m\circ\mu^*_{n,j})(a,W)\left(\mu^*_{n,j}(a,W)-\mu_0(a,W)\right)\Big\}\,\Big|\,\mathcal{D}_n\Bigg] \\
&\lesssim \mathbb{I}_{g_1,g_2}\max_{j\in[J]} P_0\left|\phi\{\mu^*_{n,j}(a,\cdot)-\mu_0(a,\cdot)\}^2\right| \\
&\lesssim \mathbb{I}_{g_1,g_2}\max_{j\in[J]} P_0\left|\phi\cdot(\mu^*_{n,j}-\mu_0)^2\right|.
\end{aligned}
$$

Combining the above bounds and recalling that $\phi=\phi_1-\phi_2$, we find

$$
\begin{aligned}
|\text{(I)}+\text{(II)}| &\lesssim \max_{j\in[J]} P_0\left|(\phi_1-\phi_2)\cdot(\mu_0-\mu_{n,j})(\pi_0-\pi_{n,j})\right| + \mathbb{I}_{g_1,g_2}\max_{j\in[J]} P_0\left|(\phi_1-\phi_2)(\mu^*_{n,j}-\mu_0)^2\right| \\
&\lesssim \|\phi_1-\phi_2\|_\infty\left\{\max_{j\in[J]}\|\mu_0-\mu_{n,j}\|\|\pi_0-\pi_{n,j}\| + \mathbb{I}_{g_1,g_2}\max_{j\in[J]}\|\mu^*_{n,j}-\mu_0\|^2\right\} \\
&\lesssim \delta_\infty\left\{\max_{j\in[J]}\|\mu_0-\mu_{n,j}\|\|\pi_0-\pi_{n,j}\| + \mathbb{I}_{g_1,g_2}\max_{j\in[J]}\|\mu^*_{n,j}-\mu_0\|^2\right\} \\
&\lesssim \delta_\infty\left\{r^*_n n^{-\gamma/(2\gamma+1)} + \mathbb{I}_{g_1,g_2}(r^*_n)^2\right\}.
\end{aligned}
$$

where the second inequality follows from Cauchy-Schwarz and Events E1-E3. The first bound of the lemma then follows. The second bound follows from an identical proof, taking $\phi_1 := \phi$ and $\phi_2 := 0$ for an arbitrary $\phi \in \mathcal{F}$ with $\|\phi\| \leq \delta$.

□

Lemma B.13. *Assume the conditions of Theorem 3.4. On event A_n, for any uniformly bounded function class $\mathcal{G}$, we have*

$$\sup_{\phi\in\mathcal{G}}\left|\overline{P}_0\left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)-\overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,,\phi)\right\}\right|\lesssim \sup_{\phi\in\mathcal{G}}\|\phi\|_\infty\,(r^*_n)^2.$$

Proof of Lemma B.13. In the following, we work on event A_n. By definition and the law of iterated expectations, for $\phi \in \mathcal{G}$, we have

$$\begin{aligned}\overline{P}_0\overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,,\phi) &= \frac{1}{J}\sum_{j\in[J]}\int\frac{1}{\pi_{n,j}(a\,|\,w)}\left\{H_{m,\mu_{n,j}}(a,w)\cdot\phi(w)\right\}\left[y-\mu^*_{n,j}(a,w)\right]dP_0(o)\\ &= \frac{1}{J}\sum_{j\in[J]}\int\frac{1}{\pi_{n,j}(a\,|\,w)}\left\{H_{m,\mu_{n,j}}(a,w)\cdot\phi(w)\right\}\left[\mu_0(a,w)-\mu^*_{n,j}(a,w)\right]dP_0(o).\end{aligned}$$

Similarly,

$$\overline{P}_0\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi) = \frac{1}{J}\sum_{j\in[J]}\int\frac{1}{\pi_{n,j}(a\,|\,w)}\left\{H_{m,\mu^*_{n,j}}(a,w)\cdot\phi(w)\right\}\left[\mu_0(a,w)-\mu^*_{n,j}(a,w)\right]dP_0(o).$$

Next, for each $j \in [J]$, recall from the proof of Lemma B.8 that

$$\left|H_{m,\mu^*_{n,j}}(a,w)-H_{m,\mu_{n,j}}(a,w)\right|\lesssim\left|\mu^*_{n,j}(a,w)-\mu_{n,j}(a,w)\right|. \tag{B.2}$$

Hence, by Cauchy-Schwarz and C1a, the above display implies

$$\begin{aligned}&\left|\overline{P}_0\left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\phi)-\overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,;\phi)\right\}\right|\\ &\lesssim \max_{j\in[J]}\left\{P_0\left|\phi\left(\mu^*_{n,j}-\mu_{n,j}\right)\left(\mu^*_{n,j}-\mu_0\right)\right|\right\}\\ &\lesssim \sup_{\phi\in\mathcal{G}}\|\phi\|_\infty\max_{j\in[J]}\left\{P_0\left|\left(\mu^*_{n,j}-\mu_{n,j}\right)\left(\mu^*_{n,j}-\mu_0\right)\right|\right\}\\ &\lesssim \sup_{\phi\in\mathcal{G}}\|\phi\|_\infty\max_{j\in[J]}\left\{P_0\left(\mu^*_{n,j}-\mu_0\right)^2+P_0\left|\left(\mu_{n,j}-\mu_0\right)\left(\mu^*_{n,j}-\mu_0\right)\right|\right\}\\ &\lesssim \sup_{\phi\in\mathcal{G}}\|\phi\|_\infty\max_{j\in[J]}\left\{\left\|\mu^*_{n,j}-\mu_0\right\|^2+\left\|\mu_{n,j}-\mu_{n,j}\right\|\left\|\mu^*_{n,j}-\mu_{n,j}\right\|\right\}\\ &\lesssim \sup_{\phi\in\mathcal{G}}\|\phi\|_\infty\,(r^*_n)^2,\end{aligned}$$

where the final inequality follows from Events E1 and E3. The result then follows after taking the supremum on the left-hand side over $\phi \in \mathcal{G}$. □

Lemma B.14. *Assume the conditions of Theorem 3.4. On event A_n, for any uniformly bounded function class $\mathcal{G}$, we have*

$$\begin{aligned}\sup_{\phi\in\mathcal{G}}\left|\overline{P}_0\left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\phi-\Pi_{k(n)}\phi)\right\}\right| &\lesssim \sup_{\phi\in\mathcal{G}}\left\|\phi-\Pi_{k(n)}\phi\right\|\max_{j\in[J]}\left\|\mu^*_{n,j}-\mu_0\right\|\\ &\lesssim \sup_{\phi\in\mathcal{G}}\left\|\phi-\Pi_{k(n)}\phi\right\|r^*_n.\end{aligned}$$

Proof of Lemma B.14. In the following, we work on event A_n. Let $\phi \in \mathcal{G}$ be arbitrary. By the law of iterated expectations, C1a, C1b, and Cauchy-Schwarz, we have

$$\begin{aligned}
&\left|\overline{P}_0\left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\phi-\Pi_{k(n)}\phi)\right\}\right| \\
&\quad\lesssim \frac{1}{J}\sum_{j=1}^{J}\left|E_0\left[\pi^{-1}_{n,j}(A\,|\,W)H_{m,\mu^*_{n,j}}(A,W)\left(\phi(W)-\Pi_{k(n)}\phi(W)\right)\left(Y-\mu^*_{n,j}(A,W)\right)\,|\,\mathcal{D}_n\right]\right| \\
&\quad\lesssim \max_{j\in[J]}\left|E_0\left[\pi^{-1}_{n,j}(A\,|\,W)H_{m,\mu^*_{n,j}}(A,W)\left(\phi(W)-\Pi_{k(n)}\phi(W)\right)\left(\mu_0(A,W)-\mu^*_{n,j}(A,W)\right)\,|\,\mathcal{D}_n\right]\right| \\
&\quad\lesssim \sup_{\phi\in\mathcal{G}}\left\|\phi-\Pi_{k(n)}\phi\right\|\max_{j\in[J]}\left\|\mu^*_{n,j}-\mu_0\right\| \\
&\quad\lesssim \sup_{\phi\in\mathcal{G}}\left\|\phi-\Pi_{k(n)}\phi\right\| r^*_n,
\end{aligned}$$

where the final inequality follows from Events E1 and E3. □

B.4 Proofs of main results

B.4.1 Proofs for results of Section 3.2

We define a loss function $(o,\theta)\mapsto L(o,\theta)$ defined on $\mathcal{O}\times\mathcal{F}$ to be γ-strongly convex [Bertsekas et al., 2003] for some $\gamma>0$ if, for all $\theta_1,\theta_2\in\mathcal{F}$ and $o\in\mathcal{O}$, the following holds:

$$L(o,\theta_1)-L(o,\theta_2)\geq \frac{d}{d\varepsilon}L(o,\theta_2+\varepsilon(\theta_1-\theta_2))\big|_{\varepsilon=0}+\frac{\gamma}{2}|\theta_1(w)-\theta_2(w)|^2.$$

In fact, if $\theta(o)\mapsto L(o,\theta)$ is twice-differentiable then the loss L is γ-strongly convex if and only if the second derivative of the map $\theta(o)\mapsto L(o,\theta)$ is positive and bounded below by $\gamma/2$ (Section 3.3. of Fawzi [2017]).

Lemma B.15. *Assuming the stated conditions in Section 3.2, the minimizer* $\theta_0 := \operatorname{argmin}_{\theta\in\overline{\mathcal{F}}}R_0(\theta)$ *exists and is unique.*

Proof. γ-strong convexity of the loss function L_{μ_0} implies, for any $\theta_1,\theta_2\in\mathcal{F}$ that

$$L_{\mu_0}(w,\theta_1)-L_{\mu_0}(w,\theta_2)\geq \frac{d}{d\varepsilon}L_{\mu_0}(w,\theta_2+\varepsilon(\theta_1-\theta_2))\big|_{\varepsilon=0}+\frac{\gamma}{2}|\theta_1(w)-\theta_2(w)|^2.$$

Taking the expectation of both sides and exchanging the order of integration and differentiation gives

$$R_0(\theta_1)-R_0(\theta_2)\geq \frac{d}{d\varepsilon}R_0(\theta_2+\varepsilon(\theta_1-\theta_2))\,|_{\,\varepsilon=0}+\frac{\gamma}{2}\left\|\theta_1-\theta_2\right\|^2_{P_0}.$$

It follows that the functional $\theta\mapsto R_0(\theta)$ is strongly convex as a mapping from $L^2(P_{0,W})$ to $\mathbb{R}$. We now show that this functional is also continuous. Note for $i=1,2$ that since g_i is continuous and μ_0 has bounded range, we have $g_i\circ\mu_0$ has uniformly bounded range. Since h_1,h_2 are Lipschitz continuous, we have that $\theta(w)\mapsto L_{\mu_0}(w,\theta(w))$ is also Lipschitz continuous. Thus, there exists some constant $L>0$ such that

$$|R_0(\theta_1)-R_0(\theta_2)|\leq LE_0\,|\theta_1(W)-\theta_2(W)|\leq L\,\|\theta_1-\theta_2\|,$$

by Cauchy-Schwarz. It follows that $\theta \mapsto R_\mu(\theta)$ is a Lipschitz continuous functional defined on $L^2(P_0)$. By Theorem 5.5 of Alexanderian [2019], a strongly convex and continuous functional defined on a closed and bounded convex subset of a Hilbert space admits a unique minimizing solution. Since $\overline{\mathcal{F}}$ is closed, bounded, and convex, we conclude that $\operatorname{argmin}_{\theta \in \overline{\mathcal{F}}} R_0(\theta)$ is nonempty and contains a unique solution.

□

Lemma B.16. *Suppose there exists a constant $\gamma > 0$ such that it is P_0-almost surely true that $\sum_{a \in \mathcal{A}} c_{a,1} \cdot g_1(\mu_0(a, W)) \neq 0$, $\ddot{h}_2(\theta(W)) \neq 0$, and*

$$\frac{\ddot{h}_1(\theta(W))}{\ddot{h}_2(\theta(W))} > -\frac{\sum_{a \in \mathcal{A}} c_{a,2} \cdot g_2(\mu_0(a, W))}{\sum_{a \in \mathcal{A}} c_{a,1} \cdot g_1(\mu_0(a, W))} + \frac{\gamma}{\ddot{h}_2(\theta(W)) \sum_{a \in \mathcal{A}} c_{a,1} \cdot g_1(\mu_0(a, W))}.$$

Then, the loss L_{μ_0} of (3.1) is strongly convex.

Proof. The stated condition implies that the second derivative of the loss with respect to $\theta(w)$ holding $w \in \mathcal{W}$ fixed satisfies:

$$\frac{d^2}{d^2\theta(w)} L_{\mu_0}(w, \theta) = \ddot{h}_1(\theta(w)) \cdot \sum_{a \in \mathcal{A}} c_{a,1}(g_1 \circ \mu_0)(a, w) + \ddot{h}_2(\theta(w)) \cdot \sum_{a \in \mathcal{A}} c_{a,2}(g_2 \circ \mu_0)(a, w) > \gamma.$$

The mean value theorem combined with an exact second order Taylor expansion of $\varepsilon \mapsto L_{\mu_0}(w, \theta_2 + \varepsilon(\theta_1 - \theta_2))$ at $\varepsilon = 0$ implies, for some $\widetilde{\varepsilon}_w \in [0, 1]$ and $\widetilde{\theta}(w) := \theta_2(w) + \widetilde{\varepsilon}_w(\theta_1(w) - \theta_2(w))$, that

$$\begin{aligned} L_{\mu_0}(w, \theta_1) - L_{\mu_0}(w, \theta_2) &= \frac{d}{d\varepsilon} L_{\mu_0}(w, \theta_2 + \varepsilon(\theta_1 - \theta_2))\big|_{\varepsilon=0} + \frac{1}{2} \frac{d^2 L_{\mu_0}(w, \theta)}{d^2\theta(w)}\Big|_{\theta(w)=\widetilde{\theta}(w)} |\theta_1(w) - \theta_2(w)|^2 \\ &\geq \frac{d}{d\varepsilon} L_{\mu_0}(w, \theta_2 + \varepsilon(\theta_1 - \theta_2))\big|_{\varepsilon=0} + \frac{\gamma}{2} |\theta_1(w) - \theta_2(w)|^2, \end{aligned}$$

as desired.

□

Proof of Theorem 3.3. We first determine the efficient influence function of parameters of the form $P \mapsto E_P\{\varphi(W)(g \circ \mu_P)(a, W)\}$ for $\varphi \in \{h_m \circ \theta : \theta \in \mathcal{F}, m \in \{1, 2\}\}$ and $g \in \{g_1, g_2\}$. Note that the population risk corresponding to (3.1) can be expressed as a linear combination of the parameters of this form.

To this end, let $(P_\varepsilon : \varepsilon \in \mathbb{R}) \subset \mathcal{M}$ be an arbitrary quadratic mean differentiable submodel with $P_\varepsilon = P$ at $\varepsilon = 0$ and score $v \in L_0^2(P)$ at $\varepsilon = 0$. Abusing notation, we compute

$$\begin{aligned} \frac{d}{d\varepsilon} E_{P_\varepsilon}\{\varphi(W)(g \circ \mu_{P_\varepsilon})(a, W)\}\Big|_{\varepsilon=0} &= \frac{d}{d\varepsilon}\left[E_{P_\varepsilon}\{\varphi(W)(g \circ \mu_P)(a, W)\}\right]\Big|_{\varepsilon=0} \\ &\quad + E_P\left\{\varphi(W) \frac{d}{d\varepsilon}(g \circ \mu_{P_\varepsilon})(a, W)\Big|_{\varepsilon=0}\right\} \\ &= E_P\{\varphi(W)(g \circ \mu_P)(a, W) v(O)\} \\ &\quad + \frac{d}{d\varepsilon}\left[E_P\{\varphi(W)(\dot{g} \circ \mu_P)(a, W)\mu_{P_\varepsilon}(a, W)\}\right]\Big|_{\varepsilon=0}. \end{aligned}$$

Note,

$$\begin{aligned}
&\frac{d}{d\varepsilon}\left[E_P\left\{\varphi(W)(\dot{g}\circ\mu_P)(a,W)\mu_{P_\varepsilon}(a,W)\right\}\right]\Big|_{\varepsilon=0}\\
&\quad=\frac{d}{d\varepsilon}\left[\int\left\{\varphi(w)(\dot{g}\circ\mu_P)(a,w)\frac{1(s=a)Y}{\pi_P(a\,|\,w)}\right\}P_\varepsilon(dy\,|\,a,w)P_{A,W}(s,w)\right]\Big|_{\varepsilon=0}\\
&\quad=E_P\left\{\varphi(W)(\dot{g}\circ\mu_P)(a,W)\frac{1(A=a)Y}{\pi_P(a\,|\,W)}v_Y(O)\right\}\\
&\quad=E_P\left\{\varphi(W)(\dot{g}\circ\mu_P)(a,W)\frac{1(A=a)(Y-\mu_P(A,W))}{\pi_P(a\,|\,W)}v(O)\right\}dP(O),
\end{aligned}$$

where $o\mapsto v_Y(o):=v(o)-E_P[v(O)\,|\,A=a,W=w]\in L_0^2(P)$ is the projection of v onto the Hilbert subspace consisting of functions that are, conditional on A,W, mean-zero functions of Y. The efficient influence function of the parameter $P\mapsto E_P\left\{\varphi(W)(g\circ\mu_P)(a,W)\right\}$ can now be read off as

$$\begin{aligned}
D:o\mapsto&\varphi(w)(g\circ\mu_P)(s,w)-E_P[\varphi(W)(g\circ\mu_P)(s,W)]\\
&+\left\{\varphi(W)(\dot{g}\circ\mu_P)(s,w)\frac{1(a=s)}{\pi_P(a\,|\,w)}(Y-\mu_P(a,w))\right\}
\end{aligned}$$

noting that $\frac{d}{d\varepsilon}E_{P_\varepsilon}\left\{\varphi(W)(g\circ\mu_{P_\varepsilon})(a,W)\right\}\Big|_{\varepsilon=0}=\langle D,v\rangle_{L^2(P)}$. The result then follows from linearity of the derivative operator, noting, for $\theta\in\mathcal{F}$, that

$$R_P(\theta)=\sum_{s\in\mathcal{A}}c_{s,2}E_P\left\{\varphi_1(W)(g_1\circ\mu_P)(s,W)\right\}+\sum_{s\in\mathcal{A}}c_{s,2}E_P\left\{\varphi_2(W)(g_2\circ\mu_P)(s,W)\right\}$$

is a linear combination of parameters of the above form, where $\varphi_m:=h_m\circ\theta$ for $m\in\{1,2\}$. □

B.4.2 Proof of Theorem 3.2

Proof. Before proceeding with the proof, we introduce some notation. Denote $\Pi_{k(n)}\mathcal{H}:=\{\Pi_{k(n)}\phi:\phi\in\mathcal{H}\}\subset\mathcal{H}_{k(n)}$, and $\mathcal{H}-\Pi_{k(n)}\mathcal{H}:=\{\phi-\Pi_{k(n)}\phi:\phi\in\mathcal{H}\}$. Define the sieve approximation rate $\rho_{n,\infty}:=n^{-1/2}+\{\log k(n)\}^\nu k(n)^{-\rho}$ where $\nu,\rho>0$ are the exponents of C3. We note, by Lemma B.1, it holds that $\sup_{\phi\in\mathcal{H}}\|\phi-\Pi_{k(n)}\phi\|_\infty\lesssim\rho_{n,\infty}$. We recall that $\mathbb{I}_{g_1,g_2}$ denotes the indicator that takes the value 0 if the functions g_1 and g_2 in (3.1) are the identity function and 1 otherwise. We will assume, without loss of generality, that event A_n and, therefore, events E1-E4, occur.

Proof strategy. To establish the result of the theorem, we proceed as follows. First, we establish that the debiasing term is asymptotically negligible in that

$$\|\overline{P}_n\Delta_{\pi_{n,\diamond},\mu^*_{n,\diamond}}\|_{\ell^\infty(\mathcal{F})}:=\sup_{\theta\in\mathcal{F}}|\overline{P}_n\Delta_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot;\theta)|=\mathcal{O}_p(n^{-1/2}).$$

Noting that $\overline{P}_n\Delta_{\pi_{n,\diamond},\mu^*_{n,\diamond}}=R_{n,k(n)}-R_{n,\pi_{n,\diamond},\mu^*_{n,\diamond}}$, this implies that $\|R_{n,k(n)}-R_{n,\pi_{n,\diamond},\mu^*_{n,\diamond}}\|_{\ell^\infty(\mathcal{F})}=\mathcal{O}_p(n^{-1/2})$. Consequently, the EP-learner risk estimator $R_{n,k(n)}$ is asymptotically equivalent to the one-step risk estimator $R_{n,\pi_{n,\diamond},\mu^*_{n,\diamond}}$. Afterwards, using the previous result, we show that the EP-learner risk estimator $R_{n,k(n)}$ satisfies $\|R_{n,k(n)}-R_{n,0}\|_{\ell^\infty(\mathcal{F})}=\mathcal{O}_p(n^{-1/2})$ and is,

thus, asymptotically equivalent to the oracle-efficient one-step risk estimator $R_{n,0}$. To do so, it suffices, by the triangle inequality, to establish that $\|R_{n,\pi_n,\mu_n^*} - R_{n,0}\|_{\ell^\infty(\mathcal{F})} = \mathcal{O}_p(n^{-1/2})$. The desired weak convergence result then follows from weak convergence of $R_{n,0}$ to R_0 in $\ell^\infty(\mathcal{F})$ and a functional form of Slutsky's lemma.

Controlling the debiasing term. We begin by controlling the debiasing remainder term $\|\overline{P}_n \Delta_{\pi_{n,\diamond},\mu^*_{n,\diamond}}\|_{\ell^\infty(\mathcal{F})}$. By the definition of $\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot,\phi)$ for $\phi \in \mathcal{H}$, we have that

$$\sup_{\theta\in\mathcal{F}} \left|\overline{P}_n \Delta_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\theta)\right| \lesssim \max_{m\in\{1,2\}} \sup_{\phi\in\mathcal{H}} \left|\overline{P}_n \overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)\right|.$$

Hence, it suffices to bound, for each $m \in \{1,2\}$, the term $\sup_{\phi\in\mathcal{H}} \left|\overline{P}_n \overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)\right|$. Noting that $\phi \mapsto \overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)$ is linear as a mapping in ϕ, the triangle inequality gives

$$\begin{aligned}\sup_{\phi\in\mathcal{H}} \left|\overline{P}_n \overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)\right| &\le \sup_{\phi\in\mathcal{H}} \left|\overline{P}_n \overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi)\right| \\ &\quad + \sup_{\phi\in\mathcal{H}} \left|\overline{P}_n \left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\phi) - \overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi)\right\}\right|.\end{aligned}$$

We denote the two terms on the right-hand side of the above display as:

$$\begin{aligned}\text{(I)} &:= \sup_{\phi\in\mathcal{H}} \left|\overline{P}_n \overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi)\right|; \\ \text{(II)} &:= \sup_{\phi\in\mathcal{H}} \left|\overline{P}_n \left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\phi) - \overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi)\right\}\right|.\end{aligned}$$

We bound each of the above terms in turn. To bound (I), recall that the first-order equations characterizing the debiased outcome regression estimators $(\mu^*_{n,j} : j \in [J])$ of Algorithm 4 imply that

$$\sup_{\phi\in\mathcal{H}} \left|\overline{P}_n \overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi)\right| = 0.$$

Next, note that $\overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}} = \overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}$ if $\mathbb{I}_{g_1,g_2} = 0$. Hence, in view of the previous display, term (I) is zero in the case where $\mathbb{I}_{g_1,g_2} = 0$. If $\mathbb{I}_{g_1,g_2} = 1$, we can further expand term (I) as follows:

$$\begin{aligned}\text{(I)} &= \mathbb{I}_{g_1,g_2} \sup_{\phi\in\mathcal{H}} \left|\overline{P}_n \overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi) + \overline{P}_n \left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi) - \overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi)\right\}\right| \\ &\le \mathbb{I}_{g_1,g_2} \sup_{\phi\in\mathcal{H}} \left|\overline{P}_n \overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi)\right| \\ &\quad + \mathbb{I}_{g_1,g_2} \sup_{\phi\in\mathcal{H}} \left|\overline{P}_n \left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi) - \overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi)\right\}\right| \\ &\le \text{(Ia)} + \text{(Ib)} + \text{(Ic)},\end{aligned}$$

where, using that $\overline{P}_n = \overline{P}_0 + (\overline{P}_n - \overline{P}_0)$, we define:

$$\begin{aligned}\text{(Ia)} &:= \mathbb{I}_{g_1,g_2} \sup_{\phi\in\mathcal{H}} \left|\overline{P}_n \overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi)\right|; \\ \text{(Ib)} &:= \mathbb{I}_{g_1,g_2} \sup_{\phi\in\mathcal{H}} \left|\overline{P}_0 \left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi) - \overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi)\right\}\right|; \\ \text{(Ic)} &:= \mathbb{I}_{g_1,g_2} \sup_{\phi\in\mathcal{H}} \left|(\overline{P}_n - \overline{P}_0) \left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi) - \overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,;\Pi_{k(n)}\phi)\right\}\right|.\end{aligned}$$

The first order derivative equations solved by $(\mu^*_{n,j} : j \in [J])$ imply that term (Ia) is zero. We will bound term (Ib) using the Cauchy-Schwarz inequality and term (Ic) using empirical process techniques. Note, by C3 and C4, we have $\sup_{\phi\in\mathcal{H}} \left\|\Pi_{k(n)}\phi\right\|_\infty \leq \sup_{\phi\in\mathcal{H}} \|\phi\|_\infty + \sup_{\phi\in\mathcal{H}} \left\|\phi - \Pi_{k(n)}\phi\right\|_\infty = O(1) + \rho_{n,\infty} = O(1)$, since $\rho_{n,\infty} = o(1)$. Hence, we have $\Pi_{k(n)}\mathcal{H}$ is a uniformly bounded function class. By Lemma B.13, uniform boundedness of $\Pi_{k(n)}\mathcal{H}$, and C2c, we have

$$\text{(Ib)} \lesssim \mathbb{I}_{g_1,g_2}(r^*_n)^2.$$

Next, applying Lemma B.8 with the function class $\mathcal{G} := \Pi_{k(n)}\mathcal{H}$, and the entropy integral bound of Lemma B.3, we find that

$$\text{(Ic)} = \mathbb{I}_{g_1,g_2}\mathcal{O}_p\left(n^{-1/2}r^*_n\mathcal{J}_\infty\left(1/r^*_n, \mathcal{F}\right) + r^*_n\sqrt{k(n)\log n/n}\right). \tag{B.3}$$

Turning to term (II), we introduce the following bound:

$$\begin{aligned}
\text{(II)} &= \sup_{\phi\in\mathcal{H}}\left|\overline{P}_n\left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\phi - \Pi_{k(n)}\phi)\right\}\right| \\
&\leq \sup_{\phi\in\mathcal{H}}\left|\overline{P}_0\left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\phi - \Pi_{k(n)}\phi)\right\}\right| + \sup_{\phi\in\mathcal{H}}\left|(\overline{P}_n - \overline{P}_0)\left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\phi - \Pi_{k(n)}\phi)\right\}\right| \\
&\leq \text{(IIa)} + \text{(IIb)} + \text{(IIc)},
\end{aligned}$$

where we define:

$$\begin{aligned}
\text{(IIa)} &:= \sup_{\phi\in\mathcal{H}}\left|\overline{P}_0\left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\phi - \Pi_{k(n)}\phi)\right\}\right|; \\
\text{(IIb)} &:= \sup_{\phi\in\mathcal{H}}\left|(\overline{P}_n - \overline{P}_0)\left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu_0}(\cdot\,;\phi - \Pi_{k(n)}\phi)\right\}\right|; \\
\text{(IIc)} &:= \sup_{\phi\in\mathcal{H}}\left|(\overline{P}_n - \overline{P}_0)\left\{\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\phi - \Pi_{k(n)}\phi) - \Delta_{\pi_{n,\diamond},\mu_0}(\cdot\,;\phi - \Pi_{k(n)}\phi)\right\}\right|.
\end{aligned}$$

By Lemma B.14, term (IIa) satisfies

$$\text{(IIa)} = \mathcal{O}_p\left(\sup_{\phi\in\mathcal{H}}\left\|\phi - \Pi_{k(n)}\phi\right\| r^*_n\right) = \mathcal{O}_p\left(\rho_{n,\infty}r^*_n\right).$$

Next, applying Lemma B.3, Lemma B.10 with $\mathcal{G} := \mathcal{H} - \Pi_{k(n)}\mathcal{H}$, and Markov's inequality, we find that

$$\text{(IIb)} = \mathcal{O}_p\left(n^{-1/2}\mathcal{J}_\infty(\max\{n^{-1/2}, \rho_{n,\infty}\}, \mathcal{F})\right).$$

Similarly, by Lemma B.9, C2c, and Markov's inequality, we have that

$$\text{(IIc)} = \mathcal{O}_p\left(n^{-1/2}r^*_n\mathcal{J}_\infty\left((\rho_{n,\infty} + n^{-1/2})/r^*_n, \mathcal{F}\right) + r^*_n\rho_{n,\infty}\sqrt{k(n)\log n/n}\right).$$

Combining all the bounds for (I) and (II), we obtain

$$\begin{aligned}
\text{(I)} &= \mathbb{I}_{g_1,g_2}\mathcal{O}_p\left(n^{-2\beta/(2\beta+1)} + k(n)\log n/n + n^{-1/2}r^*_n\mathcal{J}_\infty\left(1/r^*_n, \mathcal{F}\right)\right); \\
\text{(II)} &= \mathcal{O}_p\Big(\left\{\rho_{n,\infty} + \rho_{n,\infty}\sqrt{k(n)\log n/n}\right\}\left\{n^{-\beta/(2\beta+1)} + \sqrt{k(n)\log n/n}\right\} \\
&\quad + n^{-1/2}r^*_n\mathcal{J}_\infty\left(\max\{\rho_{n,\infty}, n^{-1/2}\}/r^*_n, \mathcal{F}\right) + n^{-1/2}\mathcal{J}_\infty\left(\max\{\rho_{n,\infty}, n^{-1/2}\}, \mathcal{F}\right)\Big).
\end{aligned}$$

We will now show that (I) + (II) $= o_p(n^{-1/2})$ and, so, $\|\overline{P}_n\Delta_{\pi_{n,\diamond},\mu^*_{n,\diamond}}\|_{\ell^\infty(\mathcal{F})} = \mathcal{O}_p(n^{-1/2})$.

To do so, we first show that the following entropy integral remainders are negligible:

$$n^{-1/2} r_n^* \mathcal{J}_\infty \left(1/r_n^*, \mathcal{F}\right) + n^{-1/2} r_n^* \mathcal{J}_\infty \left(\max\{\rho_{n,\infty}, n^{-1/2}\}/r_n^*, \mathcal{F}\right) + n^{-1/2} \mathcal{J}_\infty \left(\max\{\rho_{n,\infty}, n^{-1/2}\}, \mathcal{F}\right) = \mathcal{O}_p(n^{-1/2}). \tag{B.4}$$

Note $r_n^* = o(1)$ by C2b, and the upper bound on the growth rate $k(n)$. Also, by C3 and the lower bound on the sieve growth rate $k(n)$, we have $\rho_{n,\infty} = o(1)$. Thus, by C4, each term in the above display is $o_p(n^{-1/2})$.

In view of the above displays, to show that (I) is $o_p(n^{-1/2})$, it remains to show that

$$\mathbb{I}_{g_1,g_2} O_p\left(n^{-2\beta/(2\beta+1)} + k(n)\log n/n\right) = \mathcal{O}_p(n^{-1/2}).$$

By C2b, the outcome regression rate exponent satisfies $\beta > 1/2$ when $\mathbb{I}_{g_1,g_2} = 1$ and, thus, $n^{-2\beta/(2\beta+1)} = o_P(n^{-1/2})$. Moreover, since the sieve growth rate satisfies $k(n) = o(\sqrt{n}/\log n)$, we have $k(n)\log n/n = \mathcal{O}_p(n^{-1/2})$. Finally, to show that (II) is $o_p(n^{-1/2})$, it remains to show that, under C3, $\rho_{n,\infty}\sqrt{k(n)\log n/n} = \mathcal{O}_p(n^{-1/2})$ and $\rho_{n,\infty} n^{-2\beta/(2\beta+1)} = \mathcal{O}_p(n^{-1/2})$. The former statement holds since $\rho > 1/2$ and, thus, $\rho_{n,\infty}\sqrt{k(n)\log n} = \{\log k(n)\}^\nu k(n)^{-\rho+1/2}(\log n) = \mathcal{O}_p(1)$ by C3. The latter statement holds since $k(n)$ satisfies:

$$\begin{aligned}
&\rho_{n,\infty} n^{-\beta/(2\beta+1)} = \{\log k(n)\}^\nu k(n)^{-\rho}(1/n)^{\beta/(2\beta+1)} = o(n^{-1/2}); \\
\iff &\{\log k(n)\}^{2\nu} k(n)^{-2\rho} = o(n^{-1} n^{2\beta/(2\beta+1)}) = o(n^{-1/(2\beta+1)}); \\
\iff &k(n) = \{\log k(n)\}^{\nu/\rho}\omega(n^{(1/2\rho)/(2\beta+1)}),
\end{aligned}$$

where the final expression is true by assumption. We conclude that $\|\overline{P}_n \Delta_{\pi_{n,\diamond},\mu^*_{n,\diamond}}\|_{\ell^\infty(\mathcal{F})} = \mathcal{O}_p(n^{-1/2})$, which establishes the first statement of the theorem.

Establishing oracle-efficiency. Next, we establish asymptotic equivalence with the oracle efficient one-step estimator $(R_{n,0}(\theta) : \theta \in \mathcal{F})$ and, consequently, weak convergence of the EP-learner risk estimator. We begin by demonstrating weak convergence of the oracle one-step risk estimator $(R_{n,0}(\theta) : \theta \in \mathcal{F})$. First, the class $\mathcal{F}$ is Donsker since, by C4, it has finite uniform entropy integral (Theorem 2.8.3 of van der Vaart and Wellner 1996). Moreover, by C1a and Theorem 3.3, the risk functional $P \mapsto R_P(\theta)$ is pathwise differentiable with efficient influence function $D_0(\cdot\,;\theta)$ for each $\theta \in \mathcal{F}$. Note that the map $\theta \mapsto R_{n,0}(\theta) = R_0(\theta) + P_n D_0(\cdot\,;\theta)$ is pointwise asymptotically linear with θ-specific influence function being the efficient influence function $D_0(\cdot\,;\theta)$. By C1a and Lipschitz-continuity of (h_1, h_2) in (3.3), the class $(D_0(\cdot\,;\theta) : \theta \in \mathcal{F})$ is a Lipschitz transformation of the Donsker class $\mathcal{M}$, and is, thus, also Donsker by preservation of the Donsker property (Theorem 2.10.6 of van der Vaart and Wellner 1996). Hence, by the functional central limit theorem for Donsker classes (Theorem 2.8.2 of van der Vaart and Wellner 1996), the process $(\sqrt{n}\{R_{n,0}(\theta) - R_0(\theta)\} : \theta \in \mathcal{F})$ converges weakly in $\ell^\infty(\mathcal{F})$ to a tight mean-zero Gaussian process with the claimed covariance structure. Next, we establish weak convergence of EP-learner risk estimator $(\sqrt{n}\{R_{n,k(n)}(\theta) - R_0(\theta)\} : \theta \in \mathcal{F})$ to the same limit in $\ell^\infty(\mathcal{F})$. By Slutsky's lemma for weak convergence in Banach spaces (Lemma 1.10.2 of van der Vaart and Wellner 1996), it suffices to show that $\|R_{n,k(n)} - R_{n,0}\|_{\ell^\infty(\mathcal{F})} = \sup_{\theta\in\mathcal{F}} \left|R_{n,k(n)}(\theta) - R_{n,0}(\theta)\right| = \mathcal{O}_p(n^{-1/2})$, so that $R_{n,k(n)}$ is asymptotically equivalent to the oracle one-step risk estimator.

By the triangle inequality, this weak convergence result follows so long as

$$\sup_{\theta\in\mathcal{F}}\left|R_{n,k(n)}(\theta)-R_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\theta)\right|+\sup_{\theta\in\mathcal{F}}\left|R_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\theta)-R_{n,0}(\theta)\right|=\mathcal{O}_p(n^{-1/2}).$$

Observe that the first term $\sup_{\theta\in\mathcal{F}}\left|R_{n,k(n)}(\theta)-R_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\theta)\right|$ equals the uniform debiasing term $\sup_{\theta\in\mathcal{F}}\left|\overline{P}_n\Delta_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,\theta)\right|$. Hence, by the first part of this proof, this term is $o_p(n^{-1/2})$. We now turn to the second term. Note the oracle loss is unbiased in that $\overline{P}_0L_{\pi_0,\mu_0}(\cdot\,,\theta)=\overline{P}_0L_{\mu_0}(\theta)$. Therefore, we have the expansion:

$$\begin{aligned}
&\sup_{\theta\in\mathcal{F}}\left|R_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\theta)-R_{n,0}(\theta)\right|\\
&\quad=\sup_{\theta\in\mathcal{F}}\left|\overline{P}_n\left\{L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\theta)-L_{\pi_0,\mu_0}(\cdot\,,\theta)\right\}\right|\\
&\quad=\sup_{\theta\in\mathcal{F}}\left|(\overline{P}_n-\overline{P}_0)\left\{L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\theta)-L_{\pi_0,\mu_0}(\cdot\,,\theta)\right\}+\overline{P}_0\left\{L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\theta)-L_{\pi_0,\mu_0}(\cdot\,,\theta)\right\}\right|\\
&\quad\leq \text{(I)}+\text{(II)},
\end{aligned}$$

where we define the terms:

$$\begin{aligned}
\text{(I)}&:=\sup_{\theta\in\mathcal{F}}\left|(\overline{P}_n-\overline{P}_0)\left\{L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,;\theta)-L_{\pi_0,\mu_0}(\cdot\,,\theta)\right\}\right|;\\
\text{(II)}&:=\sup_{\theta\in\mathcal{F}}\left|\overline{P}_0\left\{L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\theta)-L_{\mu_0}(\cdot\,,\theta)\right\}\right|.
\end{aligned}$$

By Lemma B.11, boundedness of $\mathcal{F}$, and Markov's inequality, we have term (I) satisfies

$$\text{(I)}=\mathcal{O}_p\left(n^{-1/2}s^*_n\mathcal{J}_\infty\left(1/s^*_n,\mathcal{F}\right)\right)+\mathcal{O}_p\left(r^*_n\sqrt{k(n)\log n/n}\right).$$

By Condition C2, C2a, and C2b and the upper bound on the growth rate $k(n)$, we have s^*_n and r^*_n are $o(1)$. Hence, by C4, the first term on the right-hand side of the above display is $o_p(n^{-1/2})$. Observe that the second term on the right-hand side is $o_p(n^{-1/2})$ so long as both $k(n)\log n/n=\mathcal{O}_p(n^{-1/2})$ and $n^{-\beta/(2\beta+1)}\sqrt{k(n)\log n/n}=\mathcal{O}_p(n^{-1/2})$. The first rate holds since $k(n)=o(\sqrt{n}/\log n)$. The second rate holds since, by C2,

$$n^{-\beta/(2\beta+1)}\sqrt{k(n)\log n/n}=n^{-(1/2)/(2\beta+1)}\sqrt{k(n)\log n}=o(n^{-1/2}),$$

where, for the final inequality, we use that $k(n)=o(n^{2\beta/(2\beta+1)}/\log n$ by assumption. Putting it all together, we conclude that term (I) is $o_p(n^{-1/2})$. To bound term (II), note by Lemma B.12, C2c, and Event E3 that

$$|\text{(II)}|\leq\sup_{\theta\in\mathcal{F}}\left|\overline{P}_0\left\{L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\theta)-L_{\mu_0}(\cdot\,,\theta)\right\}\right|\lesssim r^*_n n^{-\gamma/(2\gamma+1)}+\mathbb{I}_{g_1,g_2}(r^*_n)^2.$$

By C2 we have $\beta>1/(4\gamma)$ and, therefore, $n^{-\beta/(2\beta+1)}n^{-\gamma/(2\gamma+1)}=\mathcal{O}_p(n^{-1/2})$. By C2 and the upper bound on the sieve growth rate $k(n)=o(n^{2\gamma/(2\gamma+1)})$, we have $\sqrt{k(n)\log n/n}\cdot n^{-\gamma/(2\gamma+1)}=\mathcal{O}_p(n^{-1/2})$. Thus, the first term of the previous display is $o_p(n^{-1/2})$. The second term on the right-hand side of the above display is $o_p(n^{-1/2})$ by C2 and the upper bound on the growth rate $k(n)=o(n^{1/2}/\log n$. Combining all our bounds, we conclude that $\sup_{\theta\in\mathcal{F}}\left|R_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\theta)-R_{n,0}(\theta)\right|=\mathcal{O}_p(n^{-1/2})$ from which the weak convergence result follows. This completes the proof.

□

B.4.3 Additional technical lemmas for Theorems 3.3 and 3.4

We recall the oracle empirical risk minimizer $\theta_{n,0} = \operatorname{argmin}_{\theta\in\mathcal{F}} R_{n,0}(\theta)$ of the oracle efficient one-step risk estimator $R_{n,0}$. To establish oracle-efficency of the ERM-based EP-learner with respect to the oracle learner $\theta_{n,0}$, we require the following additional technical lemmas. Together, these lemmas establish that the excess risk $R_{n,0}(\theta) - R_{n,0}(\theta_{n,0})$ at the oracle minimizer $\theta_{n,0}$ can be lower bounded by the quadratic term $\gamma\|\theta - \theta_{n,0}\|^2$, up to a negligible empirical process remainder. Importantly, this result holds even if $R_{n,0}$ is nonconvex.

Lemma B.17 (Approximate strong convexity of oracle-efficient risk estimator)**.** *Under the setup of Section 3.2.2, there exists some constant $\gamma > 0$ such that for any $\theta \in \mathcal{F}$ we have*

$$R_{n,0}(\theta) - R_{n,0}(\theta_{n,0}) \geq \gamma\,\|\theta - \theta_{n,0}\|^2 + \frac{1}{2}\sum_{m\in\{1,2\}} (\overline{P}_n - \overline{P}_0)\left\{w_{0,m} h_m''(\widetilde{\theta}_n)(\theta - \theta_{n,0})^2\right\},$$

where $\widetilde{\theta}_n \in \mathcal{F}$ is some random function.

Proof of Lemma B.17. Since $\mathcal{F}$ is convex, we have, for any $\theta \in \mathcal{F}$, that $\theta_{n,0} + (\theta - \theta_{n,0})\varepsilon \in \mathcal{F}$ for all $\varepsilon \in [0,1]$. Moreover,

$$\frac{d}{d\varepsilon} R_{n,0}(\theta_{n,0} + (\theta - \theta_{n,0})\varepsilon)\big|_{\varepsilon=0} = \lim_{\varepsilon\downarrow 0} \frac{R_{n,0}(\theta_{n,0} + (\theta - \theta_{n,0})\varepsilon) - R_{n,0}(\theta_{n,0})}{\varepsilon} \geq 0,$$

since $\theta_{n,0}$ minimizes $R_{n,0}$ over $\mathcal{F}$. Recall that the functions h_1, h_2 in the definition of $R_{n,0}$ are assumed twice continuously differentiable. By the mean value theorem and a second-order Taylor expansion of $\varepsilon \mapsto R_{n,0}(\theta_{n,0} + (\theta - \theta_{n,0})\varepsilon)$ around $\varepsilon = 0$, there exists some random $\varepsilon_{n,0} \in [0,1]$ such that

$$\begin{aligned}
&R_{n,0}(\theta) - R_{n,0}(\theta_{n,0}) \\
&\quad= \frac{d}{d\varepsilon} R_{n,0}(\theta_{n,0} + (\theta - \theta_{n,0})\varepsilon)\big|_{\varepsilon=0} + \frac{1}{2}\frac{d^2}{d\varepsilon^2} R_{n,0}(\theta_{n,0} + (\theta - \theta_{n,0})\varepsilon)\big|_{\varepsilon=\varepsilon_{n,0}} \\
&\quad\geq \frac{1}{2}\frac{d^2}{d\varepsilon^2} R_{n,0}(\theta_{n,0} + (\theta - \theta_{n,0})\varepsilon)\big|_{\varepsilon=\varepsilon_{n,0}}.
\end{aligned}$$

We can write the loss $L_{\pi_0,\mu_0}(\theta, \cdot)$ corresponding to the risk $R_{n,0}(\theta) = \overline{P}_n L_{\pi_0,\mu_0}(\theta, \cdot)$ as

$$L_{\pi_0,\mu_0}(\theta, o) = \sum_{m\in\{1,2\}} w_{0,m}(y, a, w)(h_m \circ \theta)(w),$$

where, for $m \in \{1,2\}$, we define the weight function,

$$o \mapsto w_{0,m}(y, a, w) := \left\{\sum_{s\in\mathcal{A}} c_{s,m}(g_m \circ \mu_0)(s, w)\right\} + \frac{1}{\pi_0(a, w)} H_{m,\mu_0}(a, w)\,\{y - \mu_0(a, w)\},$$

which is unbiased in that $E[w_{0,m}(Y, A, W) \,|\, W = w] = \sum_{s\in\mathcal{A}} c_{s,m}(g_m \circ \mu)(s, w)$. Now, denote

$\theta_\varepsilon := \theta_{n,0} + \varepsilon(\theta - \theta_{n,0})$ and observe that

$$\frac{d^2}{d\varepsilon^2} R_{n,0}(\theta_{n,0} + (\theta - \theta_{n,0})\varepsilon)\big|_{\varepsilon=\varepsilon_{n,0}} = \sum_{m\in\{1,2\}} \overline{P}_n \left\{ w_{0,m} h_m''(\theta_{\varepsilon_{n,0}})(\theta - \theta_{n,0})^2 \right\},$$

where $\theta_{\varepsilon_{n,0}} \in \mathcal{F}$. We also have that

$$\begin{aligned}\sum_{m\in\{1,2\}} \overline{P}_n \left\{ w_{0,m} h_m''(\theta_{\varepsilon_{n,0}})(\theta - \theta_{n,0})^2 \right\} &= \sum_{m\in\{1,2\}} \overline{P}_0 \left\{ w_{0,m} h_m''(\theta_{\varepsilon_{n,0}})(\theta - \theta_{n,0})^2 \right\} \\ &\quad + \sum_{m\in\{1,2\}} (\overline{P}_n - \overline{P}_0) \left\{ w_{0,m} h_m''(\theta_{\varepsilon_{n,0}})(\theta - \theta_{n,0})^2 \right\}.\end{aligned}$$

We will show that the first term on the right-hand side can be lowered bounded by $\gamma \|\theta - \theta_{n,0}\|^2$ for some $\gamma > 0$. To this end, using that $E_0[w_{0,m}(Y, A, W) \,|\, W = w] = \sum_{s\in\mathcal{A}} c_{s,m}(g_m \circ \mu)(s, w)$ for $m \in \{1, 2\}$, we find

$$\begin{aligned}&\sum_{m\in\{1,2\}} \overline{P}_0 \left\{ w_{0,m} h_m''(\theta_{\varepsilon_{n,0}})(\theta - \theta_{n,0})^2 \right\} \\ &= \sum_{m\in\{1,2\}} E_0 \left\{ h_m''(\theta_{\varepsilon_{n,0}}(W))(\theta(W) - \theta_{n,0}(W))^2 \sum_{s\in\mathcal{A}} c_{s,m}(g_m \circ \mu)(s, W) \right\} \\ &= E_0 \left\{ (\theta(W) - \theta_{n,0}(W))^2 \ddot{\ell}_\theta(W)\big|_{\theta=\theta_{\varepsilon_{n,0}}} \right\},\end{aligned}$$

where $w \mapsto \ddot{\ell}_\theta(w) := \frac{d^2 L_{\mu_0}(\theta, w)}{d\theta(w)^2}$. We assumed in Section 3.2.2 that the loss L_{μ_0} corresponding with the risk $R_0(\theta)$ is γ-strongly convex (see Appendix B.4.1). Hence, we have that there exists some $\gamma > 0$ such that $\ddot{\ell}_\theta(W) \geq \gamma$ almost surely and, therefore, the previous display implies that

$$\sum_{m\in\{1,2\}} \overline{P}_0 \left\{ w_{0,m} h_m''(\theta_{\varepsilon_{n,0}})(\theta - \theta_{n,0})^2 \right\} \geq \gamma \|\theta - \theta_{n,0}\|^2.$$

Putting it all together, we find that

$$\begin{aligned}R_{n,0}(\theta) - R_{n,0}(\theta_{n,0}) &\geq \frac{1}{2} \sum_{m\in\{1,2\}} \overline{P}_n \left\{ w_{0,m} h_m''(\theta_{\varepsilon_{n,0}})(\theta - \theta_{n,0})^2 \right\} \\ &\geq \frac{\gamma}{2} \|\theta - \theta_{n,0}\|^2 + \frac{1}{2} \sum_{m\in\{1,2\}} (\overline{P}_n - \overline{P}_0) \left\{ w_{0,m} h_m''(\theta_{\varepsilon_{n,0}})(\theta - \theta_{n,0})^2 \right\},\end{aligned}$$

where, by convexity of $\mathcal{F}$, we have $\widetilde{\theta}_n := \theta_{\varepsilon_{n,0}} \in \mathcal{F}$ almost surely.

□

Lemma B.18. *For any random function $\widetilde{\theta}_n \in \mathcal{F}$, $\delta > 0$, and $m \in \{1, 2\}$, we have*

$$\begin{aligned}&E_0^n \left\{ \sup_{\theta_3,\theta_2,\theta_1\in\mathcal{F}:\|\theta_2-\theta_1\|\leq\delta} \left| (\overline{P}_n - \overline{P}_0) \left\{ w_{0,m} h_m''(\theta_3)(\theta_2 - \theta_1)^2 \right\} \right| \right\} \\ &\lesssim n^{-1/2} \mathcal{J}_\infty(\delta^{[2-1/(2\alpha)]} + n^{-1/2}, \mathcal{F}).\end{aligned}$$

Proof of Lemma B.18 . Consider the function class $\mathcal{G} := \left\{ w_{0,m} h_m''(\theta_3)(\theta_2 - \theta_1)^2 : \theta_1, \theta_2, \theta_3 \in \mathcal{F} \right\}$.

Since h_m'' is continuous and $\mathcal{F}$ is uniformly bounded, we have that $h_m''(\theta_3)$ is uniformly bounded over all $\theta_3 \in \mathcal{F}$. Moreover, $w_{0,m}$ is uniformly bounded by Conditions C1a and C1b. It follows that $\mathcal{G}$ is a uniformly bounded function class, and, for each $\theta_1, \theta_2, \theta_3 \in \mathcal{F}$, that $\left\|w_{0,m} h_m''(\theta_3)(\theta_2 - \theta_1)^2\right\| \leq M \left\|(\theta_2 - \theta_1)^2\right\|$ for some $M > 0$. Note $(\theta_1, \theta_2, \theta_3) \mapsto w_{0,m} h_m''(\theta_3)(\theta_2 - \theta_1)^2$ is a Lipschitz-continuous map, since we use that h_m'' is Lipschitz continuous (see Section 3.2.2). Thus, Lemma B.2 implies that $\mathcal{J}_\infty(\delta, \mathcal{G}) \lesssim \mathcal{J}_\infty(\delta, \mathcal{F})$. By C5, we also have that

$$\begin{aligned}\left\|(\theta_2 - \theta_1)^2\right\| &\leq \|\theta_2 - \theta_1\|_\infty \|\theta_2 - \theta_1\| \\ &\leq \|\theta_2 - \theta_1\|^{[2-1/(2\alpha)]}.\end{aligned}$$

Additionally, an immediate application of Lemma B.4 gives

$$\begin{aligned}&E_0^n \left\{ \sup_{\theta_3, \theta_2, \theta_1 \in \mathcal{F}: \|\theta_2 - \theta_1\| \leq \delta} \left| (\overline{P}_n - \overline{P}_0) \left\{ w_{0,m} h_m''(\theta_3)(\theta_2 - \theta_1)^2 \right\} \right| \right\} \\ &\leq E_0^n \sup_{g \in \mathcal{G}: \|g\| \leq M\delta^{[2-1/(2\alpha)]}} \left| (\overline{P}_n - \overline{P}_0) g \right| \\ &\lesssim n^{-1/2} \mathcal{J}_\infty(\delta^{[2-1/(2\alpha)]} + n^{-1/2}, \mathcal{F}).\end{aligned}$$

□

B.4.4 Proofs of Theorems in Section 3.5.2

Proof of Theorem 3.3. Since $\mathcal{F}$ is convex, we have, for any $\theta \in \mathcal{F}$, that $\theta_0 + (\theta - \theta_0)\varepsilon \in \mathcal{F}$ for all $\varepsilon \in [0, 1]$. Moreover, since θ_0 minimizes R_0 over $\mathcal{F}$, it holds that

$$\frac{d}{d\varepsilon} R_0(\theta_0 + (\theta - \theta_0)\varepsilon)\big|_{\varepsilon=0} = \lim_{\varepsilon \downarrow 0} \frac{R_0(\theta_0 + (\theta - \theta_0)\varepsilon) - R_0(\theta_0)}{\varepsilon} \geq 0.$$

Using the definition of strong convexity of the risk R_0 at θ_0 provided in Section B.4.1, we have that

$$R_0(\theta_{n,0}) - R_0(\theta_0) \geq \frac{d}{d\varepsilon} R_0(\theta_0 + (\theta_{n,0} - \theta_0)\varepsilon)\big|_{\varepsilon=0} + \frac{\gamma}{2} \|\theta_{n,0} - \theta_0\|^2.$$

Thus, we have

$$\|\theta_{n,0} - \theta_0\|^2 \lesssim R_0(\theta_{n,0}) - R_0(\theta_0).$$

Next, using that $\theta_{n,0}$ is the empirical risk minimizer so that $R_{n,0}(\theta_{n,0}) - R_{n,0}(\theta_0) \leq 0$, we obtain the inequality

$$\begin{aligned}\|\theta_{n,0} - \theta_0\|^2 &\lesssim R_0(\theta_{n,0}) - R_0(\theta_0) \\ &\leq R_0(\theta_{n,0}) - R_{n,0}(\theta_{n,0}) - \{R_0(\theta_0) - R_{n,0}(\theta_{n,0})\} \\ &\lesssim (P_n - P_0) \{L_{\pi_0,\mu_0}(\theta_{n,0}) - L_{\pi_0,\mu_0}(\theta_0)\}.\end{aligned}$$

Lemma B.2 combined with Lemma B.4 implies for any $\delta > n^{-1/2}$ that

$$\phi_n(\delta) := E_0 \sup_{\theta \in \mathcal{F}: \|\theta - \theta_0\| \leq \delta:} \left| (P_n - P_0) \{L_{\pi_0,\mu_0}(\cdot, \theta) - L_{\pi_0,\mu_0}(\theta_0)\} \right| \lesssim n^{-1/2} \mathcal{J}_\infty(\delta, \mathcal{F}).$$

As in the proof of Theorem 3.4 below, to obtain the rate result, we apply Theorem 3.2.5 of van der Vaart and Wellner [1996] where, for $\theta \in \mathcal{F}$, we take $\Theta := \mathcal{F}$, $d(\theta, \theta_0) := \|\theta - \theta_0\|$, $\mathbb{M}(\theta) := -R_0(\theta)$, and $\mathbb{M}_n(\theta) := -R_{n,0}(\theta)$. We then find that $\|\theta_{n,0} - \theta_0\| = \mathcal{O}_p(\delta_{n,0})$ for any $\delta_{n,0} > n^{-1/2}$ satisfying $\phi_n(\delta_{n,0}) \leq \delta_{n,0}^2$. In view of the previous display, $\delta_{n,0} = \min\left\{\delta > n^{-1/2} : \mathcal{J}_\infty(\delta, \mathcal{F}) \leq \sqrt{n}\delta^2\right\}$ is one such choice. The result $\delta_{n,0} = O(n^{-\alpha/(2\alpha+1)})$ then follows from C4. □

Proof of Theorem 3.4. For ease of notation, we denote the EP-learner empirical risk minimizer by $\theta_n := \theta^*_{n,k(n)}$. We also denote the quantity we wish to obtain a rate for by $\delta_n := \|\theta_n - \theta_{n,0}\|$. Since θ_n minimizes the EP-learner risk $R_{n,k(n)}(\theta)$ over $\theta \in \mathcal{F}$, we have $R_{n,k(n)}(\theta_n) - R_{n,k(n)}(\theta_{n,0}) \leq 0$. Using this, we obtain the following upper bound for the excess risk:

$$\begin{aligned} R_{n,0}(\theta_n) - R_{n,0}(\theta_{n,0}) &= \{R_{n,0}(\theta_n) - R_{n,0}(\theta_{n,0})\} - \{R_{n,k(n)}(\theta_n) - R_{n,k(n)}(\theta_{n,0})\} \\ &\quad + \{R_{n,k(n)}(\theta_n) - R_{n,k(n)}(\theta_{n,0})\} \\ &\leq R_{n,0}(\theta_n) - R_{n,k(n)}(\theta_n) - R_{n,0}(\theta_{n,0}) + R_{n,k(n)}(\theta_{n,0}) \\ &\leq \phi_{n,1}(\delta_n), \end{aligned}$$

where, for $\delta > 0$, we define

$$\phi_{n,1}(\delta) := \sup_{\theta_1, \theta_2 \in \mathcal{F}: \|\theta_1 - \theta_2\| \leq \delta} \left|R_{n,0}(\theta_2) - R_{n,k(n)}(\theta_2) - R_{n,0}(\theta_1) + R_{n,k(n)}(\theta_1)\right|$$

In addition, by Lemma B.17, we have the following lower bound for the excess risk:

$$\delta_n^2 \lesssim R_{n,0}(\theta_n) - R_{n,0}(\theta_{n,0}) + \phi_{n,2}(\delta_n),$$

where, for $\delta > 0$, we define

$$\phi_{n,2}(\delta) := \sum_{m \in \{1,2\}} \sup_{\theta_1, \theta_2, \theta_3 \in \mathcal{F}: \|\theta_2 - \theta_1\| \leq \delta} \left|(\overline{P}_n - \overline{P}_0)\left\{w_{0,m} h''_m(\theta_3)(\theta_2 - \theta_1)^2\right\}\right|.$$

Finally, combining the lower and upper bounds for the excess risk and letting $\phi_n(\delta) := \phi_{n,1}(\delta) + \phi_{n,2}(\delta)$, we obtain the following inequality for δ_n:

$$\delta_n^2 \lesssim \phi_n(\delta_n).$$

We now use the above inequality to obtain a rate of convergence for δ_n. Recall that A_n is the event on which E1-E4 hold, where the constant $M > 0$ is chosen so that A_n occurs, for all n large enough, with probability at least $1 - \varepsilon$, where $\varepsilon > 0$ is arbitrary. We first bound $E_0^n\{\mathbb{I}_{A_n}\phi_n(\delta)\} = E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\} + E_0^n\{\mathbb{I}_{A_n}\phi_{n,2}(\delta)\}$ for a fixed deterministic $\delta > 0$. It follows directly from Lemma B.18 that

$$E_0^n\{\mathbb{I}_{A_n}\phi_{n,2}(\delta)\} \lesssim n^{-1/2}\mathcal{J}_\infty(\delta^{[2-1/(2\alpha)]} + n^{-1/2}, \mathcal{F}).$$

To bound $E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\}$, let $\alpha > 0$, $\rho > 1/2$, and $\nu \geq 0$ satisfy C3 and C5 and let $\rho_{n,\infty} := \{\log k(n)\}^\nu k(n)^{-\rho}$. By Lemma B.1, we have $\sup_{\phi \in \mathcal{H}} \left\|\phi - \Pi_{k(n)}\phi\right\|_\infty = O(\rho_{n,\infty})$. Lemma B.19 given at the end of this proof establishes, for any $\delta \geq n^{-1/2}$, that $E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\}$ satisfies the following bound:

$$E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\} \lesssim \delta^{1-1/(2\alpha)}\left\{r_n^* n^{-\gamma/(2\gamma+1)} + r_n^*\sqrt{k(n)\log n/n} + o(n^{-1/2})\right\}$$

$$+ \quad r_n^* \min\{\delta, \{\log k(n)\}^{\nu} k(n)^{-\rho}\} + n^{-1/2} \left[\{\log k(n)\}^{2\nu} \{k(n)^{-\rho}\}\right]^{1-1/(2\alpha)}.$$

For the moment, we assume the following bound holds and proceed with the rate of convergence proof. We begin by further bounding the expectation of $E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\}$. First, by C4, note

$$n^{-1/2} \left[\{\log k(n)\}^{2\nu} \{k(n)^{-\rho}\}\right]^{1-1/(2\alpha)} \lesssim \{\log k(n)\}^{2\nu} n^{-1/2} k(n)^{-\rho[1-1/(2\alpha)]}.$$

Next, observe that $k(n)\log n/n = o(n^{-1/2})$ since $k(n) = o(\sqrt{n}/\log n)$. Moreover, since $k(n) = o(n^{2\beta/(2\beta+1)}/\log n)$, we have $n^{-\beta/(2\beta+1)}\sqrt{k(n)\log n/n} = o(n^{-1/2})$ by C2. Hence, it holds that

$$r_n^* \sqrt{k(n)\log n/n} \le n^{-\beta/(2\beta+1)} \sqrt{k(n)\log n/n} + k(n)\log n/n = o(n^{-1/2}).$$

Similarly, since $k(n) = o(n^{2\gamma/(2\gamma+1)}/\log n)$, we have $n^{-\gamma/(2\gamma+1)}\sqrt{k(n)\log n/n} = o(n^{-1/2})$ by C2. Again, by C2, it holds that $\beta > 1/(4\gamma)$ and, thus, $n^{-\beta/(2\beta+1)} n^{-\gamma/(2\gamma+1)} = o(n^{-1/2})$. Thus, it holds that $r_n^* n^{-\gamma/(2\gamma+1)} = o(n^{-1/2})$. Combining these bounds, we find

$$\begin{aligned} E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\} \lesssim\ & o(n^{-1/2})\delta^{1-1/(2\alpha)} + r_n^* \min\{\delta, \{\log k(n)\}^{\nu} k(n)^{-\rho}\} \\ & + \{\log k(n)\}^{2\nu} \cdot n^{-1/2} k(n)^{-\rho[1-1/(2\alpha)]}. \end{aligned}$$

Using this bound for $E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\}$ and recalling that $E_0^n\{\mathbb{I}_{A_n}\phi_{n,2}(\delta)\} \lesssim n^{-1/2}\mathcal{J}_\infty(\delta^{[2-1/(2\alpha)]} + n^{-1/2}, \mathcal{F})$, it follows, for any $\delta > n^{-1/2}$, that

$$\begin{aligned} E_0^n\{\mathbb{I}_{A_n}\phi_n(\delta)\} \lesssim\ & o(n^{-1/2})\delta^{1-1/(2\alpha)} + r_n^* \min\{\delta, \{\log k(n)\}^{\nu} k(n)^{-\rho}\} \\ & + \{\log k(n)\}^{2\nu} \cdot n^{-1/2} k(n)^{-\rho[1-1/(2\alpha)]} + n^{-1/2}\mathcal{J}_\infty(\delta^{[2-1/(2\alpha)]} + n^{-1/2}, \mathcal{F}). \end{aligned} \tag{B.5}$$

To obtain the rate result, we apply the proof of Theorem 3.2.5 of van der Vaart and Wellner [1996] on the event A_n where, for $\theta \in \mathcal{F}$, we take $\Theta := \mathcal{F}$, $d(\theta, \theta_0) := \|\theta - \theta_0\|$, $\mathbb{M}(\theta) := -R_0(\theta)$, and $\mathbb{M}_n(\theta) := -R_{n,k(n)}(\theta)$. Specifically, for $\widetilde{\delta}_n := \inf\{\delta \ge n^{-1/2} : E_0^n\{\mathbb{I}_{A_n}\phi_n(\delta)\} \lesssim \delta^2\}$, the proof of Theorem 3.2.5 applied on the event A_n establishes that, for every $\varepsilon' > 0$, we can find a fixed constant $C > 0$ such that

$$P(A_n \cap \{\|\theta_n - \theta_{n,0}\| \ge C\widetilde{\delta}_n\}) \le \varepsilon'.$$

Hence,

$$\begin{aligned} P(\|\theta_n - \theta_{n,0}\| \ge C\widetilde{\delta}_n) &\le P(A_n \cap \{\|\theta_n - \theta_{n,0}\| \ge C\widetilde{\delta}_n\}) + P(A_n^c \cap \{\|\theta_n - \theta_{n,0}\| \ge C\widetilde{\delta}_n\}) \\ &\le \varepsilon' + P(A_n^c) \le \varepsilon' + \varepsilon, \end{aligned}$$

for all n large enough. Since $\varepsilon > 0$ is arbitrary, we then conclude that the EP-learner satisfies $\|\theta_n - \theta_{n,0}\| = \mathcal{O}_p(\widetilde{\delta}_n)$.

Next, we determine an upper bound for $\widetilde{\delta}_n$. By the definition of $\widetilde{\delta}_n$, it holds that $\widetilde{\delta}_n \lesssim \delta$ for any $\delta > n^{-1/2}$ such that $E_0^n\{\mathbb{I}_{A_n}\phi_n(\delta)\} \lesssim \delta^2$. In view of (B.5) and the entropy bound of C4, the inequality $E_0^n\{\mathbb{I}_{A_n}\phi_n(\delta)\} \lesssim \delta^2$ is satisfied, up to a constant, by any $\delta > n^{-1/2}$ such that

$$\delta^2 \gtrsim o(n^{-1/2})\delta^{1-1/(2\alpha)} + r_n^* \min\{\delta, \{\log k(n)\}^{\nu} k(n)^{-\rho}\} + \{\log k(n)\}^{\nu} \cdot n^{-1/2} k(n)^{-\rho[1-1/(2\alpha)]}$$

$$+ n^{-1/2}\left\{\delta^{\{2-1/(2\alpha)\}\{1-1/(2\alpha)\}}\right\}.$$

We claim the above inequality is satisfied, up to a constant, by any $\delta > n^{-1/2}$ such that all of the following hold simultaneously: (1) $o(n^{-1/2})\delta^{1-1/(2\alpha)} \lesssim \delta^2$; (2) $r_n^* \min\{\delta, \{\log k(n)\}^\nu k(n)^{-\rho}\} \lesssim \delta^2$; (3) $\{\log k(n)\}^{2\nu} \cdot n^{-1/2}k(n)^{-\rho[1-1/(2\alpha)]} \lesssim \delta^2$; and (4) $n^{-1/2}\left\{\delta^{\{2-1/(2\alpha)\}\{1-1/(2\alpha)\}}\right\} \lesssim \delta^2$. Note, by monotonicity of $\frac{\delta^{1-1/(2\alpha)}}{\delta^2}$ that $\delta' > 0$ satisfying (1) implies that any $\delta > \delta'$ also satisfies (1). A similar argument establishes the same property for solutions of inequalities (2) and (3). For (4), the same property can be established upon noting that $\delta \mapsto \frac{\delta^{\{2-1/(2\alpha)\}\{1-1/(2\alpha)\}}}{\delta^2}$ is also monotone as $[2 - 1/(2\alpha)][1 - 1/(2\alpha)] < 2$. Hence, if we can find individual solutions $\delta_{n,1}, \delta_{n,2}, \delta_{n,3}, \delta_{n,4} > n^{-1/2}$ that respectively satisfy (1), (2), (3), and (4), then their maximum $\max\{\delta_{n,1}, \delta_{n,2}, \delta_{n,3}, \delta_{n,4}\}$ will simultaneously satisfy (1)-(4). Note the first and fourth bound are satisfied by $\delta_{n,1} = \delta_{n,4} = o(n^{-\alpha/(2\alpha+1)})$, since $\alpha > 1/2$ implies $2 - 1/(2\alpha) > 1$. Next, noting $\{\log k(n)\}^{2\nu} \lesssim \{\log n\}^{2\nu}$, the third constraint is satisfied by $\delta_{n,3} > 0$ such that $\delta_{n,3}^2 = O(\{\log n\}^{2\nu} n^{-1/2} k(n)^{-\rho[1-1/(2\alpha)]})$. Lastly, note $r_n^* \min\{\delta_{n,2}, \{\log k(n)\}^\nu k(n)^{-\rho}\} \lesssim \delta_{n,2}^2$ is satisfied by some $\delta_{n,2} = O(\min\{r_n^*, \sqrt{r_n^*\{\log n\}^\nu k(n)^{-\rho}}\})$. It follows that $\max\{\delta_{n,1}, \delta_{n,2}, \delta_{n,3}, \delta_{n,4}\}$ and, hence, also $\delta_{n,1} + \delta_{n,2} + \delta_{n,3} + \delta_{n,4}$ satisfies (1)-(4), up to a constant. Taking our upper bound for $\widetilde{\delta}_n$ as $\delta_{n,1} + \delta_{n,2} + \delta_{n,3} + \delta_{n,4}$, we conclude

$$\widetilde{\delta}_n^2 \lesssim o(n^{-2\alpha/(2\alpha+1)}) + O\left(\{\log n\}^{2\nu} n^{-1/2} k(n)^{-\rho[1-1/(2\alpha)]} + r_n^* \min\{r_n^*, \{\log k(n)\}^\nu k(n)^{-\rho}\}\right). \tag{B.6}$$

Recalling that $r_n^* = n^{-\beta/(2\beta+1)} + \sqrt{k(n)\log n/n}$ and taking the second argument in the minimum operation in (B.6), we obtain the bound:

$$\begin{aligned}
\widetilde{\delta}_n^2 &\lesssim o(n^{-2\alpha/(2\alpha+1)}) + O\left(\{\log n\}^{2\nu} n^{-1/2} k(n)^{-\rho[1-1/(2\alpha)]} + r_n^* \{\log k(n)\}^\nu k(n)^{-\rho}\right) \\
&\lesssim o(n^{-2\alpha/(2\alpha+1)}) + O\left(\{\log n\}^{2\nu}\left\{n^{-\beta/(2\beta+1)} + \sqrt{k(n)\log n/n} + n^{-1/2}k(n)^{\rho/(2\alpha))}\right\} k(n)^{-\rho}\right) \\
&\lesssim o(n^{-2\alpha/(2\alpha+1)}) + O\left(\{\log n\}^{2\nu}\left\{n^{-\beta/(2\beta+1)} + \sqrt{k(n)\log n/n} + \sqrt{k(n)^{(\rho/\alpha)}/n}\right\} k(n)^{-\rho}\right).
\end{aligned}$$

Taking the first argument in the minimum operation in (B.6), we also have the following bound:

$$\widetilde{\delta}_n^2 \lesssim o(n^{-2\alpha/(2\alpha+1)}) + O\left(n^{-2\beta/(2\beta+1)} + k(n)\log n/n + \{\log n\}^{2\nu} n^{-1/2} k(n)^{-\rho[1-1/(2\alpha)]}\right).$$

Taking the square root of both sides of the above bounds gives the desired rate of the Theorem. To complete the proof, it remains to prove Lemma B.19, which we do next. □

Recall from the proof of Theorem 3.4, for $\delta > 0$, that

$$\phi_{n,1}(\delta) := \sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta} \left|R_{n,0}(\theta_2) - R_{n,k(n)}(\theta_2) - R_{n,0}(\theta_1) + R_{n,k(n)}(\theta_1)\right|.$$

Further recall that $r_n^* := n^{-\beta/(2\beta+1)} + \sqrt{k(n)\log n/n}$, $s_n^* := r_n^* + n^{-\gamma/(2\gamma+1)}$ and, moreover, $\rho_{n,\infty} := \{\log k(n)\}^\nu k(n)^{-\rho}$. The follow lemma concerns the expected value of this quantity.

Lemma B.19 (Proof of bound for $E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\}$)**.** *For any $\delta > n^{-1/2}$, it holds that*

$$E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\} \lesssim \delta^{1-1/(2\alpha)}\left\{r_n^* n^{-\gamma/(2\gamma+1)} + r_n^*\sqrt{k(n)\log n/n} + o(n^{-1/2})\right\} + r_n^*\min\{\delta, \{\log k(n)\}^{\nu}k(n)^{-\rho}\} + n^{-1/2}\left[\{\log k(n)\}^{2\nu}\{k(n)^{-\rho}\}\right]^{1-1/(2\alpha)}.$$

Proof. As shorthand, we denote $\left[F - \widetilde{F}\right][\theta_1 - \theta_2] := \left[F(\theta_1) - \widetilde{F}(\theta_1)\right] - \left[F(\theta_2) - \widetilde{F}(\theta_2)\right]$ for any two functionals $F, \widetilde{F} \in \ell^\infty(\mathcal{F})$.

To bound $E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\}$, we proceed along lines similar to the proof of Theorem 3.2. Adding and subtracting, we have, by the triangle inequality, that

$$E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\} \le E_0^n\left[\mathbb{I}_{A_n}\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta}\left|\left[R_{n,0} - R_{\pi_{n,\diamond},\mu^*_{n,\diamond}}\right][\theta_1-\theta_2]\right|\right] + E_0^n\left[\mathbb{I}_{A_n}\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta}\left|\left[R_{\pi_{n,\diamond},\mu^*_{n,\diamond}} - R_{n,k(n)}\right][\theta_1-\theta_2]\right|\right]. \tag{B.7}$$

Turning to the second term in the above bound, note that

$$\begin{aligned}
&E_0^n\left[\mathbb{I}_{A_n}\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta}\left|\left[R_{\pi_{n,\diamond},\mu^*_{n,\diamond}} - R_{n,k(n)}\right][\theta_1-\theta_2]\right|\right]\\
&\quad= E_0^n\left[\mathbb{I}_{A_n}\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta}\left|\overline{P}_n\left\{\Delta_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot,\theta_1) - \Delta_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot,\theta_2)\right\}\right|\right]\\
&\quad\le \sum_{m\in\{1,2\}} E_0^n\left[\mathbb{I}_{A_n}\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta}\left|\overline{P}_n\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot, h_m\circ\theta_1 - h_m\circ\theta_2)\right|\right],
\end{aligned}$$

where we recall, by definition, that the map $\phi \mapsto \overline{P}_n\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)$ is linear for $m \in \{1,2\}$. Hence, we have

$$E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\} \le E_0^n\left[\mathbb{I}_{A_n}\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta}\left|\left[R_{n,0} - R_{\pi_{n,\diamond},\mu^*_{n,\diamond}}\right][\theta_1-\theta_2]\right|\right] \le \text{(I)} + \text{(II)},$$

where we denote:

$$\begin{aligned}
\text{(I)} &:= E_0^n\left[\mathbb{I}_{A_n}\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta}\left|\left[R_{n,0} - R_{\pi_{n,\diamond},\mu^*_{n,\diamond}}\right][\theta_1-\theta_2]\right|\right];\\
\text{(II)} &:= \sum_{m\in\{1,2\}} E_0^n\left[\mathbb{I}_{A_n}\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\le\delta}\left|\overline{P}_n\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot, h_m\circ\theta_1 - h_m\circ\theta_2)\right|\right]\\
&= \sum_{m\in\{1,2\}} E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}^\star_\delta}\left|\overline{P}_n\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot,\phi)\right|\right],
\end{aligned}$$

where we let $\mathcal{H}^\star_\delta := \{h_m\circ\theta_1 - h_m\circ\theta_2 : \theta_1,\theta_2\in\mathcal{F}, \|\theta_1-\theta_2\|\le\delta, m\in\{1,2\}\}$.

We bound each of the above terms in turn. To bound term (I), consider the decomposition:

$$\begin{aligned}
&\left[R_{n,0} - R_{\pi_{n,\diamond},\mu^*_{n,\diamond}}\right][\theta_1-\theta_2]\\
&= \left[\overline{P}_n\left(L_{\pi_0,\mu_0} - L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}\right)\right][\theta_1-\theta_2]\\
&= \left[\overline{P}_0\left(L_{\pi_0,\mu_0} - L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}\right)\right][\theta_1-\theta_2] + \left[(\overline{P}_n - \overline{P}_0)\left(L_{\pi_0,\mu_0} - L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}\right)\right][\theta_1-\theta_2].
\end{aligned}$$

By C5, it holds that $\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\leq\delta}\|\theta_1-\theta_2\|_\infty \leq \delta^{1-1/(2\alpha)}$. Thus, applying Lemma B.12, it holds that

$$\begin{aligned}\mathbb{I}_{A_n}\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\leq\delta}&\left|\left[\overline{P}_0\left(L_{\pi_0,\mu_0}-L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}\right)\right][\theta_1-\theta_2]\right|\\ &\lesssim \sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\leq\delta}\|\theta_1-\theta_2\|_\infty\left\{r_n^* n^{-\gamma/(2\gamma+1)}+\mathbb{I}_{g_1,g_2}(r_n^*)^2+r_n^*\sqrt{k(n)\log n/n}\right\}\\ &\lesssim \delta^{1-1/(2\alpha)}\left\{r_n^* n^{-\gamma/(2\gamma+1)}+\mathbb{I}_{g_1,g_2}(r_n^*)^2+r_n^*\sqrt{k(n)\log n/n}\right\},\end{aligned}$$

where the right-hand side is deterministic. Note that $\mathbb{I}_{g_1,g_2}(r_n^*)^2 = o(n^{-1/2})$ by C2 and the constraint that $\mathbb{I}_{g_1,g_2}k(n) = \mathbb{I}_{g_1,g_2}o(n^{2\beta/(2\beta+1)}/\log n)$. Hence, the term $\delta^{1-1/(2\alpha)}\mathbb{I}_{g_1,g_2}(r_n^*)^2 = \delta^{1-1/(2\alpha)}o(n^{-1/2})$. Next, the entropy integral bound of C4 implies that $s_n^*\mathcal{J}_\infty(\delta/s_n^*,\mathcal{F}) \lesssim \delta^{1-1/(2\alpha)}(s_n^*)^{1/(2\alpha)}$. Hence, by Lemma B.11, it holds that

$$\begin{aligned}E_0^n\left[\mathbb{I}_{A_n}\sup_{\theta_1,\theta_2\in\mathcal{F}:\|\theta_1-\theta_2\|\leq\delta}\left|\left[(\overline{P}_n-\overline{P}_0)\left(L_{\pi_0,\mu_0}-L_{\pi_{n,\diamond},\mu^*_{n,\diamond}}\right)\right][\theta_1-\theta_2]\right|\right]&\\ \lesssim n^{-1/2}s_n^*\mathcal{J}_\infty(\delta/s_n^*,\mathcal{F})&\\ \lesssim \delta^{1-1/(2\alpha)}\left\{n^{-1/2}(s_n^*)^{1/(2\alpha)}\right\}.&\end{aligned}$$

Consequently, applying both of the above bounds and applying the triangle inequality, we find

$$\text{(I)} \lesssim \delta^{1-1/(2\alpha)}\left[r_n^* n^{-\gamma/(2\gamma+1)}+o(n^{-1/2})+r_n^*\sqrt{k(n)\log n/n}+n^{-1/2}(s_n^*)^{1/(2\alpha)}\right].$$

Since $r_n^* = o(1)$, we have $n^{-1/2}r_n^*\mathcal{J}_\infty(\delta/r_n^*,\mathcal{F}) \lesssim \delta^{1-1/(2\alpha)}o(n^{-1/2})$. Moreover, the fact that $\sqrt{k(n)\log n/n} \leq r_n^*$ implies that

$$\mathbb{I}_{g_1,g_2}\delta^{1-1/(2\alpha)}r_n^*\sqrt{k(n)\log n/n} \leq \mathbb{I}_{g_1,g_2}\delta^{1-1/(2\alpha)}(r_n^*)^2 = \delta^{1-1/(2\alpha)}o(n^{-1/2}).$$

We now turn to term (II). Note, by Lipschitz continuity of the fixed functions h_1 and h_2 in the definition of R_0, there exists some Lipschitz constant $L > 0$ such that $\|h_m\circ\theta_1 - h_m\circ\theta_2\|_\infty \leq L\|\theta_1-\theta_2\|_\infty$ and $\|h_m\circ\theta_1 - h_m\circ\theta_2\| \leq L\|\theta_1-\theta_2\|$ for all $\theta_1,\theta_2\in\mathcal{F}$ and $m\in\{1,2\}$. In view of this, we define the function classes:

$$\begin{aligned}\mathcal{H}_\delta &:= \left\{\phi_1-\phi_2 : \phi_1,\phi_2\in\mathcal{H};\ \|\phi_1-\phi_2\|\leq L\delta;\ \|\phi_1-\phi_2\|_\infty \leq L\delta^{1-1/(2\alpha)}\right\};\\ \mathcal{H}_{1,\delta}^{k(n)} &:= \left\{\phi-\Pi_{k(n)}\phi : \phi\in\mathcal{H}_\delta\right\};\ \mathcal{H}_{2,\delta}^{k(n)} := \left\{\Pi_{k(n)}\phi : \phi\in\mathcal{H}_\delta\right\}.\end{aligned}$$

Importantly, for $m\in\{1,2\}$, using the norm coupling of C5, it follows that $h_m\circ\theta_1 - h_m\circ\theta_2 \in \mathcal{H}_\delta$ for all $\theta_1,\theta_2\in\mathcal{F}$ with $\|\theta_1-\theta_2\|\leq\delta$. Hence, $\mathcal{H}_\delta^\star \subseteq \mathcal{H}_\delta$, which we will use to bound the supremum over $\mathcal{H}_\delta^\star$ appearing in term (II) by a supremum over $\mathcal{H}_\delta$. We will then use the function classes $\mathcal{H}_{1,\delta}^{k(n)}$ and $\mathcal{H}_{2,\delta}^{k(n)}$ to further bound the supremum over $\mathcal{H}_\delta$.

Before proceeding with the proof, we derive norm and entropy integral bounds for $\mathcal{H}_\delta$, $\mathcal{H}_{1,\delta}^{k(n)}$, and $\mathcal{H}_{2,\delta}^{k(n)}$. To this end, by C4, C3a, and Lemma B.3, we have $\mathcal{J}_\infty(\delta,\mathcal{H}_\delta) \lesssim \mathcal{J}_\infty(\delta,\mathcal{F})$ and $\mathcal{J}_\infty(\delta,\mathcal{H}_{1,\delta}^{k(n)})+\mathcal{J}_\infty(\delta,\mathcal{H}_{2,\delta}^{k(n)}) \lesssim \mathcal{J}_\infty(\{\log k(n)\}^\nu\delta,\mathcal{F})$ where $\nu\geq 0$. Moreover, since $\Pi_{k(n)}$ is an orthogonal projection, we have both $\|\Pi_{k(n)}\phi\| \leq \|\phi\| \lesssim \delta$ and $\|\phi-\Pi_{k(n)}\phi\| \leq \|\phi\| \lesssim \delta$ for any $\phi\in\mathcal{H}_\delta$. Also, since $\|\cdot\| \leq \|\cdot\|_\infty$ and by Lemma B.1, $\sup_{\phi\in\mathcal{H}_{1,\delta}^{k(n)}}\|\phi-\Pi_{k(n)}\phi\| \leq$

$\sup_{\phi\in\mathcal{H}_{1,\delta}^{k(n)}}\|\phi-\Pi_{k(n)}\phi\|_\infty\lesssim\rho_{n,\infty}$. Combining the preceding observations, $\sup_{\phi\in\mathcal{H}_{1,\delta}^{k(n)}}\|\phi\|\lesssim\min\{\rho_{n,\infty},\delta\}$ and $\sup_{\phi\in\mathcal{H}_{2,\delta}^{k(n)}}\|\phi\|\lesssim\delta$. In addition, by the triangle inequality, we have

$$\sup_{\phi\in\mathcal{H}_{2,\delta}^{k(n)}}\|\phi\|_\infty=\sup_{\phi\in\mathcal{H}_\delta}\|\Pi_{k(n)}\phi\|_\infty\leq\sup_{\phi\in\mathcal{H}_\delta}\|\phi-\Pi_{k(n)}\phi\|_\infty+\sup_{\phi\in\mathcal{H}_\delta}\|\phi\|_\infty\lesssim\rho_{n,\infty}+\delta^{1-1/(2\alpha)},$$

where we used that $\sup_{\phi\in\mathcal{H}_\delta}\|\phi\|_\infty\lesssim\delta^{1-1/(2\alpha)}$ by C5.

Proceeding with the proof, observe, by the triangle inequality, that

$$\begin{aligned}\text{(II)}&\lesssim\max_{m\in\{1,2\}}E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}_\delta}\left|\overline{P}_n\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)\right|\right]\\&\lesssim\text{(IIA)}+\text{(IIB)},\end{aligned}$$

where

$$\text{(IIA)}:=\max_{m\in\{1,2\}}E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}_{1,\delta}^{k(n)}}\left|\overline{P}_n\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)\right|\right];$$

$$\text{(IIB)}:=\max_{m\in\{1,2\}}E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}_{2,\delta}^{k(n)}}\left|\overline{P}_n\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)\right|\right].$$

To bound (II), it suffices to bound (IIA) and (IIB). To this end, term (IIA) can be further bounded as (IIA) $\leq$ (IIA1) + (IIA2) + (IIA3), where

$$\text{(IIA1)}:=\max_{m\in\{1,2\}}E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}_{1,\delta}^{k(n)}}\left|\overline{P}_0\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)\right|\right];$$

$$\text{(IIA2)}:=\max_{m\in\{1,2\}}E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}_{1,\delta}^{k(n)}}\left|\left(\overline{P}_n-\overline{P}_0\right)\left(\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu_0}(\cdot\,,\phi)\right)\right|\right];$$

$$\text{(IIA3)}:=\max_{m\in\{1,2\}}E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}_{1,\delta}^{k(n)}}\left|\left(\overline{P}_n-\overline{P}_0\right)\left(\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)-\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu_0}(\cdot\,,\phi)\right)\right|\right].$$

To bound (IIA1), we apply Lemma B.14 with $\mathcal{G}:=\mathcal{H}_{1,\delta}^{k(n)}$, the sieve approximation rate of C3, and Event E3 to obtain the bound:

$$\text{(IIA1)}\lesssim\sup_{\phi\in\mathcal{H}_{1,\delta}^{k(n)}}\|\phi-\Pi_{k(n)}\phi\|\,r_n^*\lesssim\min\{\delta,\rho_{n,\infty}\}\cdot r_n^*.$$

To bound (IIA2), we similarly apply Lemma B.10 with $\delta:=\rho_{n,\infty}+n^{-1/2}$ and the entropy bound of Lemma B.3 to obtain:

$$\begin{aligned}\text{(IIA2)}&\lesssim n^{-1/2}\mathcal{J}_\infty\left(\{\log k(n)\}^\nu\{\rho_{n,\infty}+n^{-1/2}\},\mathcal{F}\right)\\&\lesssim n^{-1/2}\left[\{\log k(n)\}^\nu\{\rho_{n,\infty}+n^{-1/2}\}\right]^{1-1/(2\alpha)}.\end{aligned}$$

To bound (IIA3), we apply Lemma B.9 and the entropy bound of Lemma B.3 to obtain:

$$\text{(IIA3)}\lesssim n^{-1/2}r_n^*\mathcal{J}_\infty\left(\left(\{\log k(n)\}^\nu\{\rho_{n,\infty}+n^{-1/2}\}\right)/r_n^*,\mathcal{F}\right)+\min\{\delta^{1-1/(2\alpha)},\rho_{n,\infty}\}r_n^*\sqrt{k(n)\log n/n}$$

$$\lesssim o\left(n^{-1/2}\left[\{\log k(n)\}^{\nu}\{\rho_{n,\infty}+n^{-1/2}\}\right]^{1-1/(2\alpha)}\right)+\min\{\delta,\rho_{n,\infty}\}r_n^*\sqrt{k(n)\log n/n},$$

where the final inequality follows from C4 and $r_n^* = o(1)$. Finally, combining the bounds for (IIA1), (IIA2), and (IIA3), we obtain the bound:

$$\text{(IIA)} \lesssim \min\{\delta,\rho_{n,\infty}\}r_n^* + n^{-1/2}\left[\{\log k(n)\}^{\nu}\{\rho_{n,\infty}+n^{-1/2}\}\right]^{1-1/(2\alpha)} + \min\{\delta,\rho_{n,\infty}\}r_n^*\sqrt{k(n)\log n/n}.$$

We now turn to (IIB). Firstly, observe that term (IIB) is zero if $\mathbb{I}_{g_1,g_2} = 0$; that is, if g_1 and g_2 in the risk definition of (3.3) equal the identity function ($x \mapsto x$). Therefore, (IIB) can be bounded as:

$$\begin{aligned}
\text{(IIB)} &= \mathbb{I}_{g_1,g_2}\max_{m\in\{1,2\}} E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}_{2,\delta}^{k(n)}}\left|\overline{P}_n\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)\right|\right]\\
&\le \mathbb{I}_{g_1,g_2}\max_{m\in\{1,2\}} E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}_{2,\delta}^{k(n)}}\left|\overline{P}_n\overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,,\phi)\right|\right]\\
&\quad+\mathbb{I}_{g_1,g_2}\max_{m\in\{1,2\}} E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}_{2,\delta}^{k(n)}}\left|(\overline{P}_n-\overline{P}_0)\left(\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\phi)-\overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,,\phi)\right)\right|\right] \qquad \text{(B.8)}\\
&\quad+\mathbb{I}_{g_1,g_2}\max_{m\in\{1,2\}} E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}_{2,\delta}^{k(n)}}\left|\overline{P}_0\left(\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)-\overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,,\phi)\right)\right|\right]. \qquad \text{(B.9)}
\end{aligned}$$

The first term on the right is zero since $\mathcal{H}_{2,\delta}^{k(n)} \subset \mathcal{H}_{k(n)}$ and the sieve-MLEs $\{\mu^*_{n,j} : j \in [J]\}$ satisfy $\overline{P}_n\overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,,\phi) = 0$ for all $\phi \in \mathcal{H}_{k(n)}$. By the triangle inequality and using that $\mathcal{H}_{2,\delta}^{k(n)} \subseteq \{\phi_1+\phi_2 : \phi_1 \in \mathcal{H}_\delta, \phi_2 \in \mathcal{H}_{1,\delta}^{k(n)}\}$, we can further upper bound (B.8) as

$$\text{(IIB)} \le \text{(IIB1)} + \text{(IIB2)} + \text{(IIB3)},$$

where we define:

$$\begin{aligned}
\text{(IIB1)} &:= \mathbb{I}_{g_1,g_2}\max_{m\in\{1,2\}} E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}_{\delta}}\left|(\overline{P}_n-\overline{P}_0)\left(\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\phi)-\overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,,\phi)\right)\right|\right]\\
\text{(IIB2)} &:= \mathbb{I}_{g_1,g_2}\max_{m\in\{1,2\}} E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}_{1,\delta}^{k(n)}}\left|(\overline{P}_n-\overline{P}_0)\left(\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\phi)-\overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,,\phi)\right)\right|\right.\\
\text{(IIB3)} &:= \mathbb{I}_{g_1,g_2}\max_{m\in\{1,2\}} E_0^n\left[\mathbb{I}_{A_n}\sup_{\phi\in\mathcal{H}_{2,\delta}^{k(n)}}\left|\overline{P}_0\left(\overline{\Delta}^{(m)}_{\pi_{n,\diamond},\mu^*_{n,\diamond}}(\cdot\,,\phi)-\overline{\Delta}^{*(m)}_{\pi_{n,\diamond},\mu_{n,\diamond}}(\cdot\,,\phi)\right)\right|\right].
\end{aligned}$$

We now bound terms (IIB1), (IIB2), and (IIB3). Notice that term (IIB2) is identical to term (IIA3), which we bounded earlier. Hence, our earlier bound implies that

$$\text{(IIB2)} \lesssim \mathbb{I}_{g_1,g_2}o\left(n^{-1/2}\left[\{\log k(n)\}^{\nu}\{\rho_{n,\infty}+n^{-1/2}\}\right]^{1-1/(2\alpha)}\right)+\mathbb{I}_{g_1,g_2}\rho_{n,\infty}r_n^*\sqrt{k(n)\log n/n}.$$

We claim that (IIB1) satisfies the bound:

$$\begin{aligned}(\text{IIB1}) &\lesssim \mathbb{I}_{g_1,g_2}\left\{n^{-1/2}r_n^*\mathcal{J}_\infty\left(\delta/r_n^*,\mathcal{F}\right)+\sup_{\phi\in\mathcal{H}_\delta}\|\phi\|_\infty r_n^*\sqrt{k(n)\log n/n}\right\}\\ &\lesssim \mathbb{I}_{g_1,g_2}\left\{n^{-1/2}r_n^*\mathcal{J}_\infty\left(\delta/r_n^*,\mathcal{F}\right)+\delta^{1-1/(2\alpha)}r_n^*\sqrt{k(n)\log n/n}\right\}.\end{aligned}$$

In the above, the first inequality follows from Lemma B.8 with $\mathcal{G}:=\mathcal{H}_\delta$ and the entropy bound $\mathcal{J}_\infty(\delta,\mathcal{H}_\delta)\lesssim\mathcal{J}_\infty(\delta,\mathcal{F})$ that we derived earlier. For the final inequality, we used that $\sup_{\phi\in\mathcal{H}_\delta}\|\phi\|_\infty\lesssim\delta^{1-1/(2\alpha)}$ by C5. Next, to bound (IIB3), note, by Lemma B.13, C3, Lemma B.3, and Event E3, that:

$$(\text{IIB3})\lesssim\mathbb{I}_{g_1,g_2}\sup_{\phi\in\mathcal{H}_{2,\delta}^{k(n)}}\|\phi\|_\infty\left\{r_n^*\right\}^2\le\mathbb{I}_{g_1,g_2}\{\delta^{1-1/(2\alpha)}+\rho_{n,\infty}\}\left\{r_n^*\right\}^2.$$

Combining the bounds for (IIB1), (IIB2), and (IIB3), we find:

$$\begin{aligned}(\text{IIB})\lesssim{}&\mathbb{I}_{g_1,g_2}n^{-1/2}\left[\{\log k(n)\}^\nu\{\rho_{n,\infty}+n^{-1/2}\}\right]^{1-1/(2\alpha)}+\mathbb{I}_{g_1,g_2}n^{-1/2}r_n^*\mathcal{J}_\infty\left(\delta/r_n^*,\mathcal{F}\right)\\&+\mathbb{I}_{g_1,g_2}\delta^{1-1/(2\alpha)}r_n^*\sqrt{k(n)\log n/n}+\mathbb{I}_{g_1,g_2}\{\delta^{1-1/(2\alpha)}+\rho_{n,\infty}\}\left\{r_n^*\right\}^2.\end{aligned}$$

Now, note that $\mathbb{I}_{g_1,g_2}(r_n^*)^2=o(n^{-1/2})$ by C2 and the constraint that $k(n)=o(n^{2\beta/(2\beta+1)}/\log n)$. Since $r_n^*=o(1)$, we have $n^{-1/2}r_n^*\mathcal{J}_\infty\left(\delta/r_n^*,\mathcal{F}\right)\lesssim\delta^{1-1/(2\alpha)}o(n^{-1/2})$. Moreover, since $\sqrt{k(n)\log n/n}\le r_n^*$, we have

$$\mathbb{I}_{g_1,g_2}\delta^{1-1/(2\alpha)}r_n^*\sqrt{k(n)\log n/n}\le\mathbb{I}_{g_1,g_2}\delta^{1-1/(2\alpha)}(r_n^*)^2=\delta^{1-1/(2\alpha)}o(n^{-1/2}).$$

Combining these bounds, we find

$$\begin{aligned}(\text{IIB})\lesssim{}&\mathbb{I}_{g_1,g_2}n^{-1/2}\left[\{\log k(n)\}^\nu\{\rho_{n,\infty}+n^{-1/2}\}\right]^{1-1/(2\alpha)}+\mathbb{I}_{g_1,g_2}o(n^{-1/2})\delta^{1-1/(2\alpha)}\\&+\mathbb{I}_{g_1,g_2}\rho_{n,\infty}\left(r_n^*\right)^2.\end{aligned}$$

Thus, combining our bounds for (IIA) and (IIB), we obtain:

$$\begin{aligned}(\text{II})\lesssim{}&r_n^*\min\{\delta,\rho_{n,\infty}\}+n^{-1/2}\left[\{\log k(n)\}^\nu\{\rho_{n,\infty}+n^{-1/2}\}\right]^{1-1/(2\alpha)}+\min\{\delta,\rho_{n,\infty}\}r_n^*\sqrt{k(n)\log n/n}\\&+\mathbb{I}_{g_1,g_2}n^{-1/2}\left[\{\log k(n)\}^\nu\{\rho_{n,\infty}+n^{-1/2}\}\right]^{1-1/(2\alpha)}+\mathbb{I}_{g_1,g_2}o(n^{-1/2})\delta^{1-1/(2\alpha)}\\&+\mathbb{I}_{g_1,g_2}\rho_{n,\infty}\left(r_n^*\right)^2.\end{aligned}$$

To bound the final term above, note, by C2 and the condition that $\mathbb{I}_{g_1,g_2}k(n)=\mathbb{I}_{g_1,g_2}o(n^{2\beta/(2\beta+1)}/\log n)$, that

$$\begin{aligned}\mathbb{I}_{g_1,g_2}\rho_{n,\infty}\left(r_n^*\right)^2&=\rho_{n,\infty}o(n^{-1/2})\\&=\left[\{\log k(n)\}^\nu\{\rho_{n,\infty}+n^{-1/2}\}\right]^{1-1/(2\alpha)}O\left(n^{-1/2}\right),\end{aligned}$$

where, for the final equality, we use that $\log k(n)=O(\log n)$ and $1-\frac{1}{2\alpha}<1$. Furthermore, for the term $\min\{\delta,\rho_{n,\infty}\}r_n^*\sqrt{k(n)\log n/n}$, we have the trivial bound $\min\{\delta,\rho_{n,\infty}\}r_n^*\sqrt{k(n)\log n/n}\le\min\{\delta,\rho_{n,\infty}\}r_n^*$, since $\sqrt{k(n)\log n/n}=o(1)$. Hence,

$$(\text{II})\lesssim r_n^*\min\{\delta,\rho_{n,\infty}\}+n^{-1/2}\left[\{\log k(n)\}^\nu\{\rho_{n,\infty}+n^{-1/2}\}\right]^{1-1/(2\alpha)}+\mathbb{I}_{g_1,g_2}o(n^{-1/2})\delta^{1-1/(2\alpha)}.$$

Here, we combined alike terms and used that $\mathbb{I}_{g_1,g_2}(r_n^*)^2 = o(n^{-1/2})$ by C2. We also used that $\rho_{n,\infty} o(n^{-1/2}) = \left[\{\log k(n)\}^\nu \{\rho_{n,\infty} + n^{-1/2}\}\right]^{1-1/(2\alpha)} O\left(n^{-1/2}\right)$, since $\log k(n) = O(\log n)$ and $1 - \frac{1}{2\alpha} < 1$.

Finally, combining our bounds for terms (I) and (II) and combining terms, we obtain the following bound:

$$\begin{aligned}
E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\} &\leq \text{(I)} + \text{(II)} \\
&\lesssim \delta^{1-1/(2\alpha)}\left[r_n^* n^{-\gamma/(2\gamma+1)} + o(n^{-1/2}) + r_n^*\sqrt{k(n)\log n/n} + n^{-1/2}(s_n^*)^{1/(2\alpha)}\right] \\
&\quad + r_n^* \min\{\delta, \rho_{n,\infty}\} + n^{-1/2}\left[\{\log k(n)\}^\nu\{\rho_{n,\infty} + n^{-1/2}\}\right]^{1-1/(2\alpha)} \\
&+ \quad \mathbb{I}_{g_1,g_2} o(n^{-1/2})\delta^{1-1/(2\alpha)} \\
&\lesssim \delta^{1-1/(2\alpha)}\left[r_n^* n^{-\gamma/(2\gamma+1)} + r_n^*\sqrt{k(n)\log n/n} + n^{-1/2}(s_n^*)^{1/(2\alpha)} + o(n^{-1/2})\right] \\
&\quad + r_n^* \min\{\delta, \rho_{n,\infty}\} + n^{-1/2}\left[\{\log k(n)\}^\nu\{\rho_{n,\infty} + n^{-1/2}\}\right]^{1-1/(2\alpha)}.
\end{aligned}$$

Recall that $\rho_{n,\infty} := \{\log k(n)\}^\nu k(n)^{-\rho} \gtrapprox n^{-1/2}$ with $\rho > 1/2$ and note, by C2 and C5, that $(s_n^*)^{1/(2\alpha)} = o(1)$. Using this, we finally obtain the desired bound:

$$\begin{aligned}
E_0^n\{\mathbb{I}_{A_n}\phi_{n,1}(\delta)\} &\lesssim \delta^{1-1/(2\alpha)}\left\{r_n^* n^{-\gamma/(2\gamma+1)} + r_n^*\sqrt{k(n)\log n/n} + o(n^{-1/2})\right\} \\
&\quad + r_n^* \min\{\delta, \{\log k(n)\}^\nu k(n)^{-\rho}\} + n^{-1/2}\left[\{\log k(n)\}^{2\nu}\{k(n)^{-\rho}\}\right]^{1-1/(2\alpha)}.
\end{aligned}$$

□

B.4.5 Proof of Theorem 3.5

Proof of Theorem 3.5. It suffices to show $\left\|\theta_{n,k(n)}^* - \theta_{n,0}\right\| = \mathcal{O}_p(n^{-\alpha/(2\alpha+1)})$ as this implies, by C6, that

$$\|\theta_{n,k(n)}^* - \theta_{n,0}\|/E_0^n\|\theta_{n,0} - \theta_0\| = \mathcal{O}_p(n^{-\alpha/(2\alpha+1)})/E_0^n\|\theta_{n,0} - \theta_0\| = \mathcal{O}_p(1),$$

as desired.

To this end, note that $n^{1/2} = o(n^{2c(\beta,\gamma)/(2c(\beta,\gamma)+1)})$ since $\beta, \gamma > 1/2$. Therefore, since $k(n) \leq o(n^{1/2}/\log n)$, we have $k(n) \leq o(n^{2c(\beta,\gamma)/(2c(\beta,\gamma)+1)}/\log n)$ satisfies the growth rate bound of Theorem 3.4. Thus, by Theorem 3.4, we have $\left\|\theta_{n,k(n)}^* - \theta_{n,0}\right\| = \mathcal{O}_p(n^{-\alpha/(2\alpha+1)}) + \mathcal{O}_p(\varepsilon_n)$, where ε_n is defined above Theorem 3.4. Hence, to show $\left\|\theta_{n,k(n)}^* - \theta_{n,0}\right\| = \mathcal{O}_p(n^{-\alpha/(2\alpha+1)})$, it suffices to show that $\varepsilon_n = \mathcal{O}_p(n^{-\alpha/(2\alpha+1)})$. In view of the minimum in the definition of ε_n and recalling that we assume $\rho = \alpha$, it further suffices to show that

$$a_n = \{\log n\}^{2\nu}\left\{n^{-\beta/(2\beta+1)} + \sqrt{k(n)\log n/n} + \sqrt{k(n)/n}\right\}k(n)^{-\alpha} = o(n^{-2\alpha/(2\alpha+1)}).$$

This holds by assumption on our growth condition on $k(n)$. For the choices of β, γ, and α required by our conditions, Lemma B.20 guarantees the existence of a sequence $k(n)$ for which this holds satisfying our conditions. We now turn to proving Lemma B.20.

□

Lemma B.20. *Suppose $\alpha > 1/2$ and $0 < \min\{\frac{1}{2}, \beta\} \le \gamma$. If $\beta > \min\{\frac{1}{2}, 2\alpha/(4\alpha^2+1)\}$, then there exists a sequence $k(n)$ such that*

$$\frac{(\log n)\, k(n)}{n^{2c(\beta,\gamma)/(2c(\beta,\gamma)+1)}} \to 0 \quad \text{and} \quad a_n\, n^{\frac{2\alpha}{2\alpha+1}} \to 0,$$

where $c(\beta,\gamma) = \min\{\gamma, \beta, 1/2\}$ and

$$a_n = \{\log n\}^{2\nu}\Big\{n^{-\beta/(2\beta+1)} + \sqrt{k(n)\log n/n} + \sqrt{k(n)/n}\Big\}\, k(n)^{-\alpha}.$$

Proof of Lemma B.20 (short version). Let $c := c(\beta,\gamma) = \min\{\gamma, \beta, 1/2\}$. Since $\gamma \ge \min\{\beta, 1/2\}$ by assumption, we have $c = \min\{\beta, 1/2\}$. Take $k(n) = n^t/(\log n)^s$ with $t \in (0,1)$ and $s > 1$.

(1) First constraint.

$$\frac{(\log n)k(n)}{n^{2c/(2c+1)}} = n^{\,t-\frac{2c}{2c+1}}(\log n)^{1-s} \to 0 \quad \Longleftrightarrow \quad t < U(c) := \frac{2c}{2c+1}.$$

(2) Second constraint.

$$a_n = \{\log n\}^{2\nu}\Big\{n^{-\beta/(2\beta+1)} + \sqrt{k(n)\log n/n} + \sqrt{k(n)/n}\Big\}k(n)^{-\alpha}.$$

Since $(\log n)^r = o(n^\delta)$ for any fixed $r \in \mathbb{R}$ and $\delta > 0$, logarithmic factors can be absorbed once the polynomial exponents in n are strictly negative. In particular, for the $\sqrt{k\log n/n}$ term, the extra $(\log n)^{1/2}$ factor is negligible compared to n^δ for some $\delta > 0$. It is therefore enough that the n-exponents are negative for

$$n^{-\beta/(2\beta+1)}k(n)^{-\alpha}n^{2\alpha/(2\alpha+1)} \quad \text{and} \quad \Big(\frac{k(n)}{n}\Big)^{1/2}k(n)^{-\alpha}n^{2\alpha/(2\alpha+1)}.$$

Writing $k(n)^{-\alpha} = n^{-\alpha t}(\log n)^{\alpha s}$ gives the two inequalities

$$-\frac{\beta}{2\beta+1} - \alpha t + \frac{2\alpha}{2\alpha+1} < 0 \quad \text{and} \quad \frac{t-1}{2} - \alpha t + \frac{2\alpha}{2\alpha+1} < 0,$$

i.e.

$$t > \frac{1}{\alpha}\Big(\frac{2\alpha}{2\alpha+1} - \frac{\beta}{2\beta+1}\Big) \quad \text{and} \quad t > \frac{1}{2\alpha+1}.$$

Hence the second constraint is equivalent to $t > L(\alpha,\beta)$ where

$$L(\alpha,\beta) := \max\Big\{\frac{1}{2\alpha+1},\ \frac{1}{\alpha}\Big(\frac{2\alpha}{2\alpha+1} - \frac{\beta}{2\beta+1}\Big)\Big\}.$$

(3) Nonemptiness of $(L(\alpha,\beta), U(c))$. Let $m := \min\{\beta, 1/2\}$ so $U(c) = U(m) = \frac{2m}{2m+1}$. It remains to show

$$\frac{1}{\alpha}\Big(\frac{2\alpha}{2\alpha+1} - \frac{\beta}{2\beta+1}\Big) \ < \ \frac{2m}{2m+1}.$$

If $\beta \le 1/2$ then $m = \beta$ and the above is equivalent to $\beta > \frac{2\alpha}{4\alpha^2+1}$. If $\beta \ge 1/2$, note that $g(\beta) := \frac{1}{\alpha}\Big(\frac{2\alpha}{2\alpha+1} - \frac{\beta}{2\beta+1}\Big)$ is decreasing in β (since $\beta \mapsto \beta/(2\beta+1)$ is increasing). Thus

$$g(\beta) \le g(1/2) = \frac{6\alpha-1}{4\alpha(2\alpha+1)} < \frac{1}{2} \quad \text{since } 8\alpha^2 - 8\alpha + 2 = 8(\alpha - \tfrac{1}{2})^2 > 0,$$

and $2m/(2m+1) = 1/2$ when $m = 1/2$. Hence in all cases $L(\alpha,\beta) < U(c)$ provided

$\beta > \min\{1/2,\, 2\alpha/(4\alpha^2+1)\}$.

Pick any $t \in \big(L(\alpha,\beta), U(c)\big)$ and, say, $s = 2$. Then both constraints hold, completing the proof. □

Recall that

$$a_n = \{\log n\}^{2\nu}\left\{n^{-\beta/(2\beta+1)} + \sqrt{k(n)\log n/n} + \sqrt{k(n)/n}\right\} k(n)^{-\alpha}$$

under the conditions of the theorem

Lemma B.21. *Suppose $\alpha > 1/2$ and $0 < \min\{\frac{1}{2}, \beta\} \le \gamma$. If $\beta > \min\{\frac{1}{2}, 2\alpha/(4\alpha^2+1)\}$, then there exists a sequence $k(n)$ such that $\frac{(\log n)\,k(n)}{n^{2c(\beta,\gamma)/(2c(\beta,\gamma)+1)}} \to 0$ and $a_n n^{\frac{2\alpha}{2\alpha+1}} \to 0$, where $c(\beta,\gamma) = \min\{\gamma, \beta, 1/2\}$*

Proof of Lemma B.21. For $k(n) = n^t/(\log n)^s$, the requirement $a_n = o(n^{-2\alpha/(2\alpha+1)})$ together with

$$\frac{(\log n)k(n)}{n^{2\beta/(2\beta+1)}} \to 0, \qquad n^{-1/2}(\log n)k(n) \to 0$$

is equivalent to

$$L(\alpha,\beta) < t < U(\beta), \qquad s > 1,$$

where

$$L(\alpha,\beta) := \max\Big\{\tfrac{1}{2\alpha+1},\, \tfrac{1}{\alpha}\Big(\tfrac{2\alpha}{2\alpha+1} - \tfrac{\beta}{2\beta+1}\Big)\Big\}, \qquad U(\beta) := \min\Big\{\tfrac{1}{2},\, \tfrac{2\beta}{2\beta+1}\Big\}.$$

We claim this interval is nonempty iff $\beta > 2\alpha/(4\alpha^2+1)$. Fix $\epsilon > 0$ small and choose σ with $L(\alpha,\beta) + \epsilon < \sigma < U(\beta)$; then

$$k(n) := \frac{n^\sigma}{(\log n)^2}$$

satisfies

$$\omega\Big(n^{\epsilon + L(\alpha,\beta)}\Big) \le k(n) \le o\big(n^{U(\beta)}/\log n\big),$$

since $\sigma < U(\beta)$ implies $n^\sigma = o(n^{U(\beta)}/\log n)$.

It remains to prove the claim. Let

$$A := \frac{1}{2\alpha+1}, \qquad B := \frac{1}{\alpha}\Big(\frac{2\alpha}{2\alpha+1} - \frac{\beta}{2\beta+1}\Big), \qquad C := \tfrac{1}{2}, \qquad D := \frac{2\beta}{2\beta+1}.$$

Then $L(\alpha,\beta) = \max\{A, B\}$ and $U(\beta) = \min\{C, D\}$. Since $A < C$ for all $\alpha > 0$, the condition $L < U$ is equivalent to $B < U$.

If $\beta \le \frac{1}{2}$, then $U = D$ and

$$B < D \iff \frac{1}{\alpha}\Big(\frac{2\alpha}{2\alpha+1} - \frac{\beta}{2\beta+1}\Big) < \frac{2\beta}{2\beta+1} \iff \beta > \frac{2\alpha}{4\alpha^2+1}.$$

If $\beta \ge \frac{1}{2}$, then $U = C = \frac{1}{2}$ and, since $A < \frac{1}{2}$, we need $B < \frac{1}{2}$. One checks that $\beta > \frac{2\alpha}{4\alpha^2+1}$ implies $B < \frac{1}{2}$, whereas if $\beta \le \frac{2\alpha}{4\alpha^2+1}$ then $B \ge D \ge U$, so $L \ge U$.

Hence $L(\alpha,\beta) < U(\beta)$ iff $\beta > \frac{2\alpha}{4\alpha^2+1}$, proving the claim and the lemma. □

B.5 Rate for debiased outcome regression nuisance

In this section, we verify that Condition C2c holds when using the squared error loss in Algorithm 4 (Method 2). Although we focus on the squared error loss here, the argument relies

primarily on the strong convexity of the loss and thus extends, with minor modifications, to more general loss functions.

The following lemma shows, under reasonable conditions, that

$$\sum_{j=1}^{J} \left\| \mu^*_{n,j} - \mu_0 \right\|_{\overline{P}_0} \lesssim \sum_{j=1}^{J} \left\| \mu_{n,j} - \mu_0 \right\|_{\overline{P}_0} + O_p\left(\sqrt{\frac{k(n) \log n}{n}} \right).$$

Assuming this bound, Condition C2c follows directly from the nuisance rate condition (Condition C2b) on the initial nuisance estimators $(\mu_{n,j} : j \in [J])$.

Let $k = k(n)$. In the following lemma, define the population analogue of β_n in Algorithm 4 as

$$\beta_0 = \operatorname*{argmin}_{\beta \in \mathbb{R}^{2k}} \sum_{j=1}^{K} E_0 \left[\frac{1}{\pi_{n,j}(A \mid W)} \left\{ Y - \mu_{n,j}(A,W) - \beta^\top \widehat{\varphi}_{k,j}(A,W) \right\}^2 \,\middle|\, \mathcal{D}_n \setminus \mathcal{D}_n^j \right].$$

Lemma B.22. *Let β_n and $\mu^*_{n,\diamond}$ be obtained from Algorithm 4 using Method 2. Assume Condition C1 and suppose that, for all n, $\max\{\|\beta_n\|_\infty, \|\beta_0\|_\infty\} < M$ almost surely. Let $k = k(n)$. Then*

$$\left\| \mu^*_{n,\diamond} - \mu_0 \right\|_{\overline{P}_0} \leq \left\| \mu_{n,\diamond} - \mu_0 \right\|_{\overline{P}_0} + O_p\left(\sqrt{\frac{k \log n}{n}} \right).$$

Proof. Let $k = k(n)$. The estimated coefficient β_n obtained from Method 2 of Algorithm 4 can be written as:

$$\beta_n = \operatorname*{argmin}_{\beta \in \mathbb{R}^{2k}} \sum_{i=1}^{n} \frac{1}{\pi_{n,j(i)}(A_i \mid W_i)} \left\{ Y_i - \mu_{n,j(i)}(A_i,W_i) - \beta^\top \widehat{\varphi}_{k,j(i)}(A_i,W_i) \right\}^2.$$

Denote the corresponding population risk minimizer by

$$\beta_0 = \operatorname*{argmin}_{\beta \in \mathbb{R}^{2k}} \sum_{j=1}^{K} E_0 \left[\frac{1}{\pi_{n,j}(A \mid W)} \left\{ Y - \mu_{n,j}(A,W) - \beta^\top \widehat{\varphi}_{k,j}(A,W) \right\}^2 \mid \mathcal{D}_n \backslash \mathcal{D}_n^j \right].$$

The first-order optimality conditions for β_n imply, for all $\beta \in \mathbb{R}^{2k}$ that

$$\sum_{i=1}^{n} \frac{1}{\pi_{n,j(i)}(A_i \mid W_i)} \{\mu_{n,j}(\beta) - \mu_{n,j}\}(A_i,W_i) \left\{ Y_i - \mu^*_{n,j(i)}(A_i,W_i) \right\} = 0,$$

where $\mu_{n,j}(\beta) := \mu_{n,j} + \beta^\top \widehat{\varphi}_{k,j}$. Using our notation, the above implies that

$$\frac{1}{J} \sum_{j=1}^{J} P_{n,j} \left[\tfrac{1}{\pi_{n,j}} \{\mu_{n,j}(\beta) - \mu_{n,j}\} \{\mathcal{I}_Y - \mu^*_{n,j}\} \right] = 0,$$

where $\mathcal{I}_Y : o \mapsto y$ denotes the y coordinate projection map. Adding and subtracting terms, we find, for all $\beta \in \mathbb{R}^{2k}$, that

$$\frac{1}{J} \sum_{j=1}^{J} P_0 \left[\tfrac{1}{\pi_{n,j}} \{\mu_{n,j}(\beta) - \mu_{n,j}\} \{\mathcal{I}_Y - \mu^*_{n,j}\} \right] = \frac{1}{J} \sum_{j=1}^{J} (P_0 - P_{n,j}) \left[\tfrac{1}{\pi_{n,j}} \{\mu_{n,j}(\beta) - \mu_{n,j}\} \{\mathcal{I}_Y - \mu^*_{n,j}\} \right].$$

Moreover, the first-order optimality conditions of β_0 imply, for all $\beta \in \mathbb{R}^{2k}$, that

$$\frac{1}{J}\sum_{j=1}^{J} P_0\Big[\tfrac{1}{\pi_{n,j}}\{\mu_{n,j}(\beta)-\mu_{n,j}\}\{\mathcal{I}_Y-\mu_{n,j}(\beta_0)\}\Big]=0.$$

Hence, combining the previous two displays, we obtain

$$\frac{1}{J}\sum_{j=1}^{J} P_0\Big[\tfrac{1}{\pi_{n,j}}\{\mu_{n,j}(\beta)-\mu_{n,j}\}\{\mu_{n,j}(\beta_0)-\mu^*_{n,j}\}\Big]=\frac{1}{J}\sum_{j=1}^{J}(P_0-P_{n,j})\Big[\tfrac{1}{\pi_{n,j}}\{\mu_{n,j}(\beta)-\mu_{n,j}\}\{\mathcal{I}_Y-\mu^*_{n,j}\}\Big].$$

Evaluating at $\beta=\beta_0$ and $\beta=\beta_n$ and subtracting the two identities yields

$$\frac{1}{J}\sum_{j=1}^{J} P_0\Big[\tfrac{1}{\pi_{n,j}}\{\mu_{n,j}(\beta_0)-\mu_{n,j}(\beta_n)\}\{\mu_{n,j}(\beta_0)-\mu^*_{n,j}\}\Big]=\frac{1}{J}\sum_{j=1}^{J}(P_0-P_{n,j})\Big[\tfrac{1}{\pi_{n,j}}\{\mu_{n,j}(\beta_0)-\mu_{n,j}(\beta_n)\}\{\mathcal{I}_Y-\mu^*_{n,j}\}\Big].$$

Since, by definition, $\mu_{n,j}(\beta_n)=\mu^*_{n,j}$, this simplifies to

$$\frac{1}{J}\sum_{j=1}^{J} P_0\Big[\tfrac{1}{\pi_{n,j}}\{\mu_{n,j}(\beta_0)-\mu^*_{n,j}\}^2\Big]=\frac{1}{J}\sum_{j=1}^{J}(P_0-P_{n,j})\Big[\tfrac{1}{\pi_{n,j}}\{\mu_{n,j}(\beta_0)-\mu^*_{n,j}\}\{\mathcal{I}_Y-\mu^*_{n,j}\}\Big].$$

Hence, using our notation and that $\frac{1}{\pi_{n,j}}<\eta^{-1}$ by C1a, we obtain the regret inequality:

$$\|\mu^*_{n,\diamond}-\mu_{n,\diamond}(\beta_0)\|^2_{\overline{P}_0}\lesssim\frac{1}{J}\sum_{j=1}^{J}(P_0-P_{n,j})\Big[\tfrac{1}{\pi_{n,j}}\{\mu_{n,j}(\beta_0)-\mu^*_{n,j}\}\{\mathcal{I}_Y-\mu^*_{n,j}\}\Big].$$

Since $\|\beta_n\|_\infty<M$ almost surely, both $\mu^*_{n,j}$ and $\mu_{n,j}(\beta_0)$ belong to the bounded function class:

$$\mathcal{H}^k_{n,j}:=\{\,\mu_{n,j}+\beta^\top\widehat{\varphi}_{k,j}:\beta\in\mathbb{R}^{2k},\|\beta\|_\infty<M\,\}.$$

Note that, since by Condition C1 both $\mu_{n,j}$ and $\widehat{\varphi}_{k,j}$ are almost surely uniformly bounded in k and j, it follows that every element of $\mathcal{H}^k_{n,j}$ is uniformly bounded by a fixed constant.

Define the random quantity $\hat{\delta}_n:=\big\|\mu^*_{n,\diamond}-\mu_{n,\diamond}(\beta_0)\big\|_{\overline{P}_0}$, which we aim to bound. To obtain a rate bound, we use that the above regret inequality implies:

$$\hat{\delta}_n^2\;\le\;\frac{1}{J}\sum_{j=1}^{J}\sup_{g_j\in\mathcal{G}^k_{n,j}:\,\|g_j\|_{P_0}\le JC\,\hat{\delta}_n}\big|(P_0-P_{n,j})\,g_j\big|,$$

where $\mathcal{G}^k_{n,j}:=\big\{\tfrac{1}{\pi_{n,j}}(\mu_{n,j}(\beta_0)-h)\,(\mathcal{I}_Y-h):h\in\mathcal{H}^k_{n,j}\big\}$, and $C>0$ is a constant such that $\big\|\tfrac{1}{\pi_{n,j}}(\mu_{n,\diamond}(\beta_0)-\mu^*_{n,\diamond})\,(\mathcal{I}_Y-\mu^*_{n,\diamond})\big\|_{\overline{P}_0}\;\le\;C\,\big\|\mu_{n,\diamond}(\beta_0)-\mu^*_{n,\diamond}\big\|_{\overline{P}_0}$, which is finite by Condition C1.

To obtain an in-probability bound on $\hat{\delta}_n$, we follow a standard ERM analysis that leverages maximal inequalities for empirical processes based on uniform entropy integrals. Conditional on the training data $\mathcal{D}_n\setminus\mathcal{D}_{n,j}$, the function class $\mathcal{G}^k_{n,j}$ has a uniform entropy integral satisfying $\mathcal{J}(\delta,\mathcal{G}^k_{n,j})\lesssim\mathcal{J}(\delta,\mathcal{H}^k_{n,j})$ by preservation of uniform entropy integrals under Lipschitz continuous transformations [van der Vaart and Wellner, 1996]. Moreover, $\mathcal{H}^k_{n,j}$ is a bounded subset of a fixed affine space of (at most) dimension $2k$ and therefore

has VC–subgraph dimension bounded by $O(k)$ [van der Vaart and Wellner, 1996]. By standard bounds on uniform entropy integrals for classes with finite VC–subgraph dimension [van der Vaart and Wellner, 1996], we have $\mathcal{J}(\delta, \mathcal{H}^k_{n,j}) \lesssim \delta\sqrt{k\log(1/\delta)}$, and, thus, $\mathcal{J}(\delta, \mathcal{G}^k_{n,j}) \lesssim \delta\sqrt{k\log(1/\delta)}$. Applying the local maximal inequality of van der Vaart and Wellner [2011] (Theorem 2.1), we have, for any fixed $\delta > \sqrt{k\log n/n}$, that

$$E^0_n\left[\frac{1}{J}\sum_{j=1}^{J}\sup_{g_j\in\mathcal{G}^k_{n,j}:\,\|g_j\|_{P_0}\le JC\,\hat{\delta}}|(P_0 - P_{n,j})\,g_j|\right] \lesssim \delta\sqrt{k\log n/n}.$$

Informally, if the above bound held with the random radius $\hat{\delta}_n$, then by Markov's inequality our regret inequality would yield

$$\hat{\delta}^2_n \lesssim \mathcal{O}_p\left(\hat{\delta}_n\sqrt{\frac{k\log n}{n}}\right),$$

from which it follows that $\hat{\delta}_n = \mathcal{O}_p\left(\sqrt{\frac{k\log n}{n}}\right)$. While the bound does not hold directly for the random $\hat{\delta}_n$, a standard peeling argument—see the proof of Theorem 3.4 and Section 3.2 in van der Vaart and Wellner [1996]—establishes that this rate is indeed correct.

Hence, we conclude that

$$\left\|\mu^*_{n,\diamond} - \mu_{n,\diamond}(\beta_0)\right\|_{\overline{P}_0} = \mathcal{O}_p\left(\sqrt{\frac{k\log n}{n}}\right).$$

Moreover, by the triangle inequality, we have that

$$\begin{aligned}\|\mu^*_{n,\diamond} - \mu_0\|_{\overline{P}_0} &\le \|\mu_{n,\diamond}(\beta_0) - \mu_0\|_{\overline{P}_0} + \|\mu^*_{n,\diamond} - \mu_{n,\diamond}(\beta_0)\|_{\overline{P}_0}\\ &\le \|\mu_{n,\diamond}(\beta_0) - \mu_0\|_{\overline{P}_0} + \mathcal{O}_p\left(\sqrt{\frac{k\log n}{n}}\right).\end{aligned}$$

Moreover, since $\mu_{n,\diamond}(\beta_0)$ minimizes the distance $\|\mu_\diamond - \mu_0\|_{\overline{P}_0}$ over $\mu_\diamond \in \mathcal{H}^k_{n,\diamond}$ and the initial estimate $\mu_{n,\diamond}$ is an element of $\mathcal{H}^k_{n,\diamond}$ (take $\beta = 0$), we have that $\left\|\mu_{n,\diamond}(\beta_0) - \mu_0\right\|_{\overline{P}_0} \le \|\mu_{n,\diamond} - \mu_0\|_{\overline{P}_0}$. Hence

$$\left\|\mu^*_{n,\diamond} - \mu_0\right\|_{\overline{P}_0} \le \|\mu_{n,\diamond} - \mu_0\|_{\overline{P}_0} + \mathcal{O}_p\left(\sqrt{\frac{k\log n}{n}}\right),$$

where k may grow with n.

□

Appendix C

PROOFS FOR CHAPTER 3

C.1 Proofs for Discussion of conditions

Proof. Fix $P \in \mathcal{P}_\delta$ and set $\Delta_P := \theta_P - \theta_0$. Since θ_P minimizes R_P, $R_P(\theta_P) \le R_P(\theta_0)$. Applying the quadratic lower bound with $\theta = \theta_P$ gives

$$R_P(\theta_P) \ge R_P(\theta_0) + \partial_\theta R_P(\theta_0)(\Delta_P) + \frac{\kappa}{2}\|\Delta_P\|_{\mathcal{H}}^2.$$

Combining the two displays yields

$$\frac{\kappa}{2}\|\Delta_P\|_{\mathcal{H}}^2 \le -\partial_\theta R_P(\theta_0)(\Delta_P).$$

Because θ_0 minimizes R_0 and R_0 is differentiable at θ_0, $\partial_\theta R_0(\theta_0) = 0$. Hence

$$\frac{\kappa}{2}\|\Delta_P\|_{\mathcal{H}}^2 \le |\{\partial_\theta R_P(\theta_0) - \partial_\theta R_0(\theta_0)\}(\Delta_P)| \le \|\partial_\theta R_P(\theta_0) - \partial_\theta R_0(\theta_0)\|_{\mathcal{H}^*}\, \|\Delta_P\|_{\mathcal{H}}.$$

If $\Delta_P = 0$, the result is immediate. Otherwise, dividing by $\|\Delta_P\|_{\mathcal{H}}$ and using assumption (ii) gives the claimed bound. □

C.2 Proofs for Section 4.3

C.2.1 Proof of Lemma 4.7

Proof of Lemma 4.7. Throughout, we work with perturbations along line segments that remain in the relevant subsets $\mathcal{H}_\mathcal{P}$ and $\mathcal{N}_\mathcal{P}$; equivalently, the pointwise arguments of l remain in the compact set $\mathcal{C}$.

We compute the first derivative of the map $t \mapsto L_0(\theta + th, \eta)$ evaluated at $t = 0$:

$$\begin{aligned}
\partial_\theta L_0(\theta,\eta)(h) &= \frac{d}{dt} L_0(\theta + th, \eta)\Big|_{t=0} \\
&= \frac{d}{dt} \int l(\theta(z) + th(z), \eta(z), z)\, P_0(dz)\Big|_{t=0} \\
&= \int \frac{d}{dt} l(\theta(z) + th(z), \eta(z), z)\Big|_{t=0} P_0(dz) \\
&= \int \partial_a l(\theta(z), \eta(z), z)^\top h(z)\, P_0(dz),
\end{aligned}$$

which shows that $\dot{\ell}_\eta(\theta, z)(h) = \partial_a l(\theta(z), \eta(z), z)^\top h(z)$. To justify the interchange of differentiation and integration, note that for each fixed z, the map

$$t \;\mapsto\; \partial_a l\big(\theta(z) + th(z), \eta(z), z\big)^\top h(z)$$

is continuous by the joint continuity of $\partial_a l$. Moreover, $\partial_a l$ and $\partial_a^2 l$ are bounded on $\mathcal{C}$ by compactness and continuity. By the mean value theorem, there exists a constant C such

that

$$\left|\partial_a l(\theta(z)+th(z),\eta(z),z)^\top h(z)\right| \;\le\; C\,\|h(z)\|_{\mathbb{R}^{d_1}} + C\,\|h(z)\|^2_{\mathbb{R}^{d_1}},$$

and the right-hand side is integrable since P_0 is a probability measure and $h \in \mathcal{H} \subseteq L^2(P_0)$. The Dominated Convergence Theorem then allows differentiation under the integral.

Next, define the remainder

$$R(h) := L_0(\theta+h,\eta) - L_0(\theta,\eta) - \partial_\theta L_0(\theta,\eta)(h).$$

A first-order Taylor expansion in a gives

$$l(\theta(z)+h(z),\eta(z),z) = l(\theta(z),\eta(z),z) + \partial_a l(\theta(z),\eta(z),z)^\top h(z) + r(z),$$

with $|r(z)| \le \frac{1}{2}\|\partial_a^2 l\|_\infty \|h(z)\|^2_{\mathbb{R}^{d_1}}$ and $\|\partial_a^2 l\|_\infty := \sup_{(a,b,z)\in\mathcal{C}} \|\partial_a^2 l(a,b,z)\| \;<\; \infty$. Integrating,

$$|R(h)| \le \tfrac{1}{2}\|\partial_a^2 l\|_\infty \int \|h(z)\|^2_{\mathbb{R}^{d_1}}\, P_0(dz) = \tfrac{1}{2}\|\partial_a^2 l\|_\infty \|h\|^2_{\mathcal{H}},$$

so $\frac{|R(h)|}{\|h\|_{\mathcal{H}}} \;\le\; \frac{1}{2}\|\partial_a^2 l\|_\infty \|h\|_{\mathcal{H}} \;\to\; 0$ as $\|h\|_{\mathcal{H}} \to 0$, verifying first-order Fréchet differentiability (Condition A3i).

For the second derivative,

$$\begin{aligned}\partial_\theta^2 L_0(\theta,\eta)(h_1,h_2) &= \frac{\partial^2}{\partial t\,\partial s} L_0\big(\theta + th_1 + sh_2, \eta\big)\big|_{t=s=0} \\ &= \int (h_1(z))^\top\, \partial_a^2 l(\theta(z),\eta(z),z)\, h_2(z)\, P_0(dz).\end{aligned}$$

Moreover, the bound

$$\left|(h_1(z))^\top\, \partial_a^2 l(\theta(z),\eta(z),z)\, h_2(z)\right| \;\le\; \|\partial_a^2 l\|_\infty\, \|h_1(z)\|_{\mathbb{R}^{d_1}}\, \|h_2(z)\|_{\mathbb{R}^{d_1}},$$

together with $h_1, h_2 \in \mathcal{H}$, allows an identical Dominated Convergence argument to interchange differentiation and integration when computing $\partial_\theta^2 L_0$.

Since, by assumption, $\partial_a^2 l(a,b,z)$ is Lipschitz in *both* a and b on $\mathcal{C}$, there exists $L > 0$ such that for any θ, η and perturbations h, g,

$$\left\|\partial_a^2 l(\theta(z)+h(z),\, \eta(z)+g(z),\, z) - \partial_a^2 l(\theta(z),\, \eta(z),\, z)\right\| \;\le\; L\big(\|h(z)\|_{\mathbb{R}^{d_1}} + \|g(z)\|_{\mathbb{R}^{d_2}}\big).$$

Hence, for any directions h_1, h_2,

$$\begin{aligned}&\left|\partial_\theta^2 L_0(\theta+h,\ \eta+g)(h_1,h_2) - \partial_\theta^2 L_0(\theta,\ \eta)(h_1,h_2)\right| \\ &\quad= \left|\int h_1(z)^\top \big[\partial_a^2 l(\theta(z)+h(z),\eta(z)+g(z),z) - \partial_a^2 l(\theta(z),\eta(z),z)\big] h_2(z)\, P_0(dz)\right| \\ &\quad\le \int \|h_1(z)\|_{\mathbb{R}^{d_1}}\, L\big(\|h(z)\|_{\mathbb{R}^{d_1}} + \|g(z)\|_{\mathbb{R}^{d_2}}\big)\, \|h_2(z)\|_{\mathbb{R}^{d_1}}\, P_0(dz) \\ &\quad\le L\,\big(\|h\|_{\mathcal{H}} + \|g\|_{\mathcal{N}}\big)\, \rho_{\mathcal{H}}(h_1)\, \|h_2\|_{\mathcal{H}},\end{aligned}$$

where the final inequality follows from the Cauchy–Schwarz inequality. In particular, by symmetry of $\partial_\theta^2 L_0(\theta,\eta)$, if $\|u\|_{\mathcal{H}} + \rho_{\mathcal{H}}(v) \le 1$, then

$$\left|\partial_\theta^2 L_0(\theta',\eta')(u,v) - \partial_\theta^2 L_0(\theta,\eta)(u,v)\right| = \left|\partial_\theta^2 L_0(\theta',\eta')(v,u) - \partial_\theta^2 L_0(\theta,\eta)(v,u)\right| \le L\,\big(\|\theta'-\theta\|_{\mathcal{H}} + \|\eta'-\eta\|_{\mathcal{N}}\big)\, \rho_{\mathcal{H}}(v)\|u\|_{\mathcal{H}},$$

and $\rho_{\mathcal{H}}(v)\|u\|_{\mathcal{H}} \le 1/4$. Absorbing this factor into the constant verifies Condition A3iii. To establish Fréchet differentiability as in A3ii, define the remainder

$$R(h_1,h_2) := \partial_\theta L_0(\theta+h_2,\eta)(h_1) - \partial_\theta L_0(\theta,\eta)(h_1) - \partial_\theta^2 L_0(\theta,\eta)(h_1,h_2).$$

By Taylor's theorem with integral remainder, we have

$$\partial_\theta L_0(\theta + h_2, \eta)(h_1) - \partial_\theta L_0(\theta, \eta)(h_1) = \int_0^1 \partial_\theta^2 L_0(\theta + t h_2, \eta)(h_1, h_2)\, dt,$$

so that

$$R(h_1, h_2) = \int_0^1 \Big[\partial_\theta^2 L_0(\theta + t h_2, \eta)(h_1, h_2) - \partial_\theta^2 L_0(\theta, \eta)(h_1, h_2)\Big] dt.$$

Using the Lipschitz property of $\partial_a^2 l$ and taking the essential supremum over z for h_1 yields

$$|R(h_1, h_2)| \le \int_0^1 L\, t\, \rho_{\mathcal{H}}(h_1)\, \|h_2\|_{\mathcal{H}}^2\, dt = \frac{L}{2}\, \rho_{\mathcal{H}}(h_1)\, \|h_2\|_{\mathcal{H}}^2.$$

In particular,

$$\sup_{\substack{h_1 \in \mathcal{H}_{\mathcal{P}} \\ \rho_{\mathcal{H}}(h_1) \le 1}} \frac{|R(h_1, h_2)|}{\|h_2\|_{\mathcal{H}}} \le \frac{L}{2}\, \|h_2\|_{\mathcal{H}} \to 0 \quad \text{as } \|h_2\|_{\mathcal{H}} \to 0,$$

which shows that the map $\theta \mapsto \partial_\theta L_0(\theta, \eta)$ is Fréchet differentiable with derivative $\partial_\theta^2 L_0(\theta, \eta)$, thereby verifying Condition A3ii.

For the cross-derivative, we have

$$\begin{aligned}\partial_\eta \partial_\theta L_0(\theta, \eta)(g, h) &= \frac{d^2}{ds\, dt}\, L_0(\theta + th, \eta + sg)\Big|_{t=s=0} \\ &= \int h(z)^\top\, \partial_a \partial_b l(\theta(z), \eta(z), z)\, g(z)\, P_0(dz).\end{aligned}$$

By the assumed Lipschitz continuity of $\partial_a \partial_b l$ in (a, b), there exists a constant $C > 0$ such that

$$\begin{aligned}&\big|h(z)^\top \big[\partial_a \partial_b l(\theta(z) + h_1(z), \eta(z) + g_1(z), z) - \partial_a \partial_b l(\theta(z), \eta(z), z)\big] g(z)\big| \\ &\qquad \le C\, \|h(z)\|_\infty\, \|g(z)\|_{\mathbb{R}^{d_2}}\, \{\|g_1(z)\|_{\mathbb{R}^{d_2}} + \|h_1(z)\|_{\mathbb{R}^{d_1}}\}.\end{aligned}$$

Hence, integrating both sides over z with respect to P_0 and applying the Cauchy–Schwarz inequality, we obtain

$$|\partial_\eta \partial_\theta L_0(\theta + h_1, \eta + g_1)(g, h) - \partial_\eta \partial_\theta L_0(\theta, \eta)(g, h)| \le L \rho_{\mathcal{H}}(h) \|g\|_{\mathcal{N}} \{\|g_1\|_{\mathcal{N}} + \|h_1\|_{\mathcal{H}}\}.$$

In particular, setting $\theta = \theta_0$ and $h_1 = 0$ gives

$$|\partial_\eta \partial_\theta L_0(\theta_0, \eta)(g, h) - \partial_\eta \partial_\theta L_0(\theta_0, \eta_0)(g, h)| \le L \rho_{\mathcal{H}}(h) \|g\|_{\mathcal{N}} \|\eta - \eta_0\|_{\mathcal{N}}.$$

If $\|g\|_{\mathcal{N}} + \rho_{\mathcal{H}}(h) \le 1$, then $\rho_{\mathcal{H}}(h)\|g\|_{\mathcal{N}} \le 1/4$, so after absorbing this factor into the constant we obtain Condition A4ii.

To establish Fréchet differentiability, define the remainder by

$$R^c(g, h) := \partial_\theta L_0(\theta_0, \eta + g)(h) - \partial_\theta L_0(\theta_0, \eta)(h) - \partial_\eta \partial_\theta L_0(\theta_0, \eta)(g, h).$$

By the integral remainder form, we write

$$\partial_\theta L_0(\theta_0, \eta+g)(h) - \partial_\theta L_0(\theta_0, \eta)(h) - \partial_\eta \partial_\theta L_0(\theta_0, \eta)(g, h) = \int_0^1 \Big[\partial_\eta \partial_\theta L_0(\theta_0, \eta + s\, g)(g, h) - \partial_\eta \partial_\theta L_0(\theta_0, \eta)(g, h)\Big] ds.$$

Thus, by the previous Lipschitz bound and integrating over s,

$$|R^c(g, h)| \le \int_0^1 C\, s\, \rho_{\mathcal{H}}(h)\, \|g\|_{\mathcal{N}}^2 ds = \frac{C}{2}\, \rho_{\mathcal{H}}(h)\, \|g\|_{\mathcal{N}}^2.$$

In particular, for all h with $\rho_{\mathcal{H}}(h) \leq 1$,

$$\frac{|R^c(g,h)|}{\|g\|_{\mathcal{N}}} \leq \frac{C}{2}\|g\|_{\mathcal{N}} \to 0 \quad \text{as } \|g\|_{\mathcal{N}} \to 0.$$

This shows that the map $\eta \mapsto \partial_\theta L_0(\theta_0,\eta)(h)$ extends to a neighborhood of $\mathcal{N}_{\mathcal{P}}$ in $\widetilde{\mathcal{N}}$ as a Fréchet differentiable map with derivative $\partial_\eta\partial_\theta L_0(\theta_0,\eta)$, thereby verifying Condition A4i, which completes the proof of A4. □

C.2.2 Technical Lemmas for functional Taylor expansions

Lemma C.1 (Quadratic expansion for target functional)**.** *Suppose that Condition A2 holds. Then, uniformly over $\theta \in \mathcal{H}_{\mathcal{P}}$ and $h \in \mathcal{H}$ such that $\theta + h \in \mathcal{H}_{\mathcal{P}}$, we have the expansion*

$$\psi_0(\theta + h) = \psi_0(\theta) + \dot{\psi}_0(\theta)(h) + Rem_\theta(h),$$

where the remainder term satisfies $|Rem_\theta(h)| \leq L\|h\|_{\mathcal{H}}^2$ for some $L < \infty$.

Proof. Fix $\theta \in \mathcal{H}_{\mathcal{P}}$ and $h \in \mathcal{H}$ such that $\theta + h \in \mathcal{H}_{\mathcal{P}}$. Since $\mathcal{H}_{\mathcal{P}}$ is convex, $\theta + th \in \mathcal{H}_{\mathcal{P}}$ for all $t \in [0,1]$. By Condition A2 and the integral form of Taylor's theorem applied to the map $t \mapsto \psi_0(\theta + t\,h)$, we have

$$\psi_0(\theta + h) - \psi_0(\theta) = \int_0^1 \frac{d}{dt}\,\psi_0(\theta + t\,h)\,dt = \int_0^1 \dot{\psi}_0(\theta + t\,h)(h)\,dt.$$

Adding and subtracting $\dot{\psi}_0(\theta)(h)$ inside the integral gives

$$\psi_0(\theta + h) = \psi_0(\theta) + \dot{\psi}_0(\theta)(h) + \int_0^1 \big[\dot{\psi}_0(\theta + t\,h) - \dot{\psi}_0(\theta)\big](h)\,dt.$$

Hence we may identify

$$\mathrm{Rem}_\theta(h) = \int_0^1 \big[\dot{\psi}_0(\theta + t\,h) - \dot{\psi}_0(\theta)\big](h)\,dt.$$

By Condition A2, the map $\theta \mapsto \dot{\psi}_0(\theta)$ is L-Lipschitz from $\|\cdot\|_{\mathcal{H}}$ to the operator norm $\|\cdot\|_{\mathcal{H}}^*$. Therefore, for each $t \in [0,1]$,

$$\big|\big[\dot{\psi}_0(\theta + t\,h) - \dot{\psi}_0(\theta)\big](h)\big| \;\leq\; L\,\|\theta + t\,h - \theta\|_{\mathcal{H}}\,\|h\|_{\mathcal{H}} = L\,t\,\|h\|_{\mathcal{H}}^2.$$

Integrating over t yields

$$\big|\mathrm{Rem}_\theta(h)\big| \leq \int_0^1 L\,t\,\|h\|_{\mathcal{H}}^2\,dt = \tfrac{1}{2}\,L\,\|h\|_{\mathcal{H}}^2 \;\leq\; L\,\|h\|_{\mathcal{H}}^2,$$

which completes the proof. □

Lemma C.2 (Quadratic expansion for derivative of risk)**.** *Suppose that Condition A3 holds. Then, uniformly over $\theta \in \mathcal{H}_{\mathcal{P}}$, $\eta \in \mathcal{N}_{\mathcal{P}}$, $h_1 \in \mathcal{H}$, and $h_2 \in \mathcal{H}$ such that $\theta + h_2 \in \mathcal{H}_{\mathcal{P}}$, we have the expansion*

$$\partial_\theta L_0(\theta + h_2, \eta)(h_1) = \partial_\theta L_0(\theta,\eta)(h_1) + \partial_\theta^2 L_0(\theta,\eta)(h_1,h_2) + Rem_{\theta,\eta}(h_1,h_2),$$

where the remainder term satisfies $|Rem_{\theta,\eta}(h_1,h_2)| \leq L\rho_{\mathcal{H}}(h_1)\|h_2\|_{\mathcal{H}}^2$ for some $L < \infty$.

Proof. Let $\theta \in \mathcal{H}_{\mathcal{P}}$, $\eta \in \mathcal{N}_{\mathcal{P}}$, $h_1 \in \mathcal{H}$, and $h_2 \in \mathcal{H}$ be such that $\theta + h_2 \in \mathcal{H}_{\mathcal{P}}$. Since $\mathcal{H}_{\mathcal{P}}$ is convex, $\theta + th_2 \in \mathcal{H}_{\mathcal{P}}$ for all $t \in [0,1]$. By Condition A3ii and the integral form of Taylor's

theorem applied to the map $t \mapsto \partial_\theta L_0(\theta + t h_2, \eta)(h_1)$, we have

$$\partial_\theta L_0(\theta + h_2, \eta)(h_1) - \partial_\theta L_0(\theta, \eta)(h_1) = \int_0^1 \partial_\theta^2 L_0(\theta + t h_2, \eta)(h_1, h_2)\, dt.$$

Thus, we may write

$$\partial_\theta L_0(\theta + h_2, \eta)(h_1) = \partial_\theta L_0(\theta, \eta)(h_1) + \partial_\theta^2 L_0(\theta, \eta)(h_1, h_2) + \mathrm{Rem}(h_1, h_2),$$

where the remainder is given by

$$\mathrm{Rem}(h_1, h_2) = \int_0^1 \left[\partial_\theta^2 L_0(\theta + t h_2, \eta)(h_1, h_2) - \partial_\theta^2 L_0(\theta, \eta)(h_1, h_2)\right] dt.$$

If $\rho_{\mathcal{H}}(h_1) = 0$ or $\|h_2\|_{\mathcal{H}} = 0$, then the desired bound is immediate. Otherwise, by symmetry of $\partial_\theta^2 L_0(\theta, \eta)$, bilinearity, and Condition A3iii applied to $\tilde{h}_1 := h_2/(2\|h_2\|_{\mathcal{H}})$ and $\tilde{h}_2 := h_1/(2\rho_{\mathcal{H}}(h_1))$, we obtain, for each $t \in [0, 1]$,

$$\begin{aligned}
&\left|\partial_\theta^2 L_0(\theta + t h_2, \eta)(h_1, h_2) - \partial_\theta^2 L_0(\theta, \eta)(h_1, h_2)\right| \\
&\quad = \left|\partial_\theta^2 L_0(\theta + t h_2, \eta)(h_2, h_1) - \partial_\theta^2 L_0(\theta, \eta)(h_2, h_1)\right| \\
&\quad = 4\, \rho_{\mathcal{H}}(h_1)\, \|h_2\|_{\mathcal{H}}^2 \left|\partial_\theta^2 L_0(\theta + t h_2, \eta)(\tilde{h}_1, \tilde{h}_2) - \partial_\theta^2 L_0(\theta, \eta)(\tilde{h}_1, \tilde{h}_2)\right| \\
&\quad \le 4L\, t\, \rho_{\mathcal{H}}(h_1)\, \|h_2\|_{\mathcal{H}}^2.
\end{aligned}$$

Absorbing the factor 4 into L and integrating over t yields

$$|\mathrm{Rem}(h_1, h_2)| \le L \rho_{\mathcal{H}}(h_1) \|h_2\|_{\mathcal{H}}^2.$$

□

Lemma C.3 (Quadratic expansion for the cross derivative of risk)**.** *Suppose that Condition A4 holds. Then, uniformly over $\eta \in \mathcal{N}_{\mathcal{P}}$, $h \in \mathcal{H}$, and $g \in \widetilde{\mathcal{N}}$ such that $\eta + g \in \mathcal{N}_{\mathcal{P}}$, we have the expansion*

$$\partial_\theta L_0(\theta_0, \eta + g)(h) = \partial_\theta L_0(\theta_0, \eta)(h) + \partial_\eta \partial_\theta L_0(\theta_0, \eta)(g, h) + Rem_\eta(g, h),$$

where $|Rem_\eta(g, h)| \le L' \rho_{\mathcal{H}}(h)\, \|g\|_{\mathcal{N}}^2$ for some constant $L' < \infty$.

Proof. Let $\eta \in \mathcal{N}_{\mathcal{P}}$, $g \in \widetilde{\mathcal{N}}$, and $h \in \mathcal{H}$ be such that $\eta + g \in \mathcal{N}_{\mathcal{P}}$. Since $\mathcal{N}_{\mathcal{P}}$ is convex, $\eta + tg \in \mathcal{N}_{\mathcal{P}}$ for all $t \in [0, 1]$. By Condition A4, the map $\eta \mapsto \partial_\theta L_0(\theta_0, \eta)$ admits an ambient open-neighborhood extension around this path, and the integral form of Taylor's theorem applied to the map $t \mapsto \partial_\theta L_0(\theta_0, \eta + t\, g)(h)$, we have

$$\partial_\theta L_0(\theta_0, \eta + g)(h) - \partial_\theta L_0(\theta_0, \eta)(h) = \int_0^1 \partial_\eta \partial_\theta L_0(\theta_0, \eta + t\, g)(g, h)\, dt.$$

Thus, we may write

$$\partial_\theta L_0(\theta_0, \eta + g)(h) = \partial_\theta L_0(\theta_0, \eta)(h) + \partial_\eta \partial_\theta L_0(\theta_0, \eta)(g, h) + \mathrm{Rem}(g, h),$$

where the remainder is given by

$$\mathrm{Rem}(g, h) = \int_0^1 \left[\partial_\eta \partial_\theta L_0(\theta_0, \eta + t\, g)(g, h) - \partial_\eta \partial_\theta L_0(\theta_0, \eta)(g, h)\right] dt.$$

If $\rho_{\mathcal{H}}(h) = 0$ or $\|g\|_{\mathcal{N}} = 0$, then the desired bound is immediate. Otherwise, by bilinearity and Condition A4ii applied to $\tilde{g} := g/(2\|g\|_{\mathcal{N}})$ and $\tilde{h} := h/(2\rho_{\mathcal{H}}(h))$, there exists a constant

$L' < \infty$ such that, for each $t \in [0,1]$,

$$\begin{aligned}
&\left|\partial_\eta \partial_\theta L_0(\theta_0, \eta + t\,g)(g,h) - \partial_\eta \partial_\theta L_0(\theta_0,\eta)(g,h)\right| \\
&\quad = 4\,\|g\|_{\mathcal{N}}\,\rho_{\mathcal{H}}(h)\left|\partial_\eta \partial_\theta L_0(\theta_0, \eta + t\,g)(\tilde{g},\tilde{h}) - \partial_\eta \partial_\theta L_0(\theta_0,\eta)(\tilde{g},\tilde{h})\right| \\
&\quad \le 4t\,L'\|g\|_{\mathcal{N}}^2\,\rho_{\mathcal{H}}(h).
\end{aligned}$$

Absorbing the factor 4 into L' and integrating over t from 0 to 1 yields

$$|\mathrm{Rem}(g,h)| \le \int_0^1 t\,L'\|g\|_{\mathcal{N}}^2\,\rho_{\mathcal{H}}(h)\,dt = \frac{L'}{2}\rho_{\mathcal{H}}(h)\,\|g\|_{\mathcal{N}}^2.$$

Hence, $|\mathrm{Rem}(g,h)| \le L'\rho_{\mathcal{H}}(h)\|g\|_{\mathcal{N}}^2$. □

C.2.3 Proof of functional von Mises expansion

Lemma C.4. *Suppose Conditions A1-A3 and A6 hold. Then, there exists a unique Hessian Riesz representer $\alpha_0 \in \overline{\mathcal{H}}$ such that the linear functional $\dot{\psi}_0(\theta_0)$ admits the representation:*

$$\dot{\psi}_0(\theta_0)(h) = \partial_\theta^2 L_0(\theta_0,\eta_0)(\alpha_0, h), \quad \forall h \in \mathcal{H}.$$

Proof. By Conditions A3ii and A6, we have the norm equivalence: there exist constants $0 < c < C < \infty$ such that $c\|\cdot\|_{\mathcal{H}} \le \sqrt{\partial_\theta^2 L_0(\theta_0,\eta_0)(\cdot,\cdot)} \le C\|\cdot\|_{\mathcal{H}}$, so that $\overline{\mathcal{H}}$ is also closed under the norm induced by the Hessian inner product $\partial_\theta^2 L_0(\theta_0,\eta_0)$. Consequently, by A2, the functional derivative $\dot{\psi}_0(\theta_0)$ defines a bounded linear functional on $\mathcal{H}$ with respect to the inner product $\partial_\theta^2 L_0(\theta_0,\eta_0)(\cdot,\cdot)$. By the Riesz representation theorem, there exists a unique Hessian Riesz representer $\alpha_0 \in \overline{\mathcal{H}}$ such that the linear functional $\dot{\psi}_0(\theta_0)$ admits the representation in (4.2). □

Proof of Theorem 4.9. Let $\theta \in \mathcal{H}_{\mathcal{P}}$, $\alpha \in \mathcal{H}$, and $\eta \in \mathcal{N}_{\mathcal{P}}$ satisfy $\max\{\rho_{\mathcal{H}}(\theta), \rho_{\mathcal{H}}(\alpha), \|\eta\|_{\mathcal{N}}\} < M'$ for some $M' < \infty$. Hereafter, we absorb M' into the constants appearing in our big-O notation.

Applying Lemma C.1 with A8 and A2 at θ_0 in the direction $h := \theta - \theta_0$, we obtain the Taylor expansion

$$\begin{aligned}
\psi_0(\theta) - \psi_0(\theta_0) - P_0\dot{\ell}_\eta(\theta)(\alpha) &= P_0\{m(\cdot,\theta) - m(\cdot,\theta_0)\} - \partial_\theta L_0(\theta,\eta)(\alpha) \\
&= \dot{\psi}_0(\theta_0)(\theta - \theta_0) - \partial_\theta L_0(\theta,\eta)(\alpha) + O\left(\|\theta - \theta_0\|_{\mathcal{H}}^2\right),
\end{aligned} \tag{C.1}$$

where the $O\left(\|\theta - \theta_0\|_{\mathcal{H}}^2\right)$ term is absent when the functional is linear. Note that by A2, A6, and the Riesz representation property of α_0 in (4.2), we have $h \mapsto \dot{\psi}_0(\theta_0)(h) = \partial_\theta^2 L_0(\theta_0,\eta_0)(\alpha_0, h)$. Thus, the first two terms on the right-hand side satisfy

$$\dot{\psi}_0(\theta_0)(\theta - \theta_0) - \partial_\theta L_0(\theta,\eta)(\alpha) = \partial_\theta^2 L_0(\theta_0,\eta_0)(\alpha_0, \theta - \theta_0) - \partial_\theta L_0(\theta,\eta)(\alpha). \tag{C.2}$$

By Lemma C.2 with A8 and A3, we have the second-order Taylor expansion:

$$\partial_\theta L_0(\theta,\eta)(\alpha) = \partial_\theta L_0(\theta_0,\eta)(\alpha) + \partial_\theta^2 L_0(\theta_0,\eta)(\theta - \theta_0, \alpha) + O(\|\theta - \theta_0\|_{\mathcal{H}}^2), \tag{C.3}$$

where the $O(\|\theta - \theta_0\|_{\mathcal{H}}^2)$ term is absent when the risk is quadratic. By Lemma C.3 with A8

and A4, we can expand the first term on the right-hand side in η:

$$\partial_\theta L_0(\theta_0,\eta)(\alpha) = \partial_\theta L_0(\theta_0,\eta_0)(\alpha) + \partial_\eta\partial_\theta L_0(\theta_0,\eta_0)(\eta-\eta_0,\alpha) + O(\|\eta-\eta_0\|_{\mathcal{N}}^2). \tag{C.4}$$

Since θ_0 minimizes the risk by A1, and the risk is smooth by A3i, the first-order optimality condition implies that $\partial_\theta L_0(\theta_0,\eta_0)(\alpha) = 0$. Moreover, $\partial_\eta\partial_\theta L_0(\theta_0,\eta_0)(\eta-\eta_0,\alpha) = 0$ by Neyman orthogonality at (θ_0,η_0) in A5. Hence, (C.4) implies that $\partial_\theta L_0(\theta_0,\eta)(\alpha) = O(\|\eta-\eta_0\|_{\mathcal{N}}^2)$. Therefore, substituting this relation into (C.3), we have

$$\partial_\theta L_0(\theta,\eta)(\alpha) = \partial_\theta^2 L_0(\theta_0,\eta)(\theta-\theta_0,\alpha) + O(\|\theta-\theta_0\|_{\mathcal{H}}^2) + O(\|\eta-\eta_0\|_{\mathcal{N}}^2).$$

Returning to (C.2), we obtain

$$\begin{aligned}\dot\psi_0(\theta_0)(\theta-\theta_0) - \partial_\theta L_0(\theta,\eta)(\alpha) = \partial_\theta^2 L_0(\theta_0,\eta_0)(\alpha_0,\theta-\theta_0) - \partial_\theta^2 L_0(\theta_0,\eta)(\theta-\theta_0,\alpha)\\ + O(\|\theta-\theta_0\|_{\mathcal{H}}^2) + O(\|\eta-\eta_0\|_{\mathcal{N}}^2).\end{aligned}$$

If $\theta=\theta_0$ or $\rho_{\mathcal{H}}(\alpha)=0$, then the left-hand side below is zero. Otherwise, by bilinearity and Condition A3iii, with $\tilde\theta := (\theta-\theta_0)/(2\|\theta-\theta_0\|_{\mathcal{H}})$ and $\tilde\alpha := \alpha/(2\rho_{\mathcal{H}}(\alpha))$, we have

$$\begin{aligned}\left|\partial_\theta^2 L_0(\theta_0,\eta)(\theta-\theta_0,\alpha) - \partial_\theta^2 L_0(\theta_0,\eta_0)(\theta-\theta_0,\alpha)\right| &= 4\|\theta-\theta_0\|_{\mathcal{H}}\rho_{\mathcal{H}}(\alpha)\left|\partial_\theta^2 L_0(\theta_0,\eta)(\tilde\theta,\tilde\alpha) - \partial_\theta^2 L_0(\theta_0,\eta_0)(\tilde\theta,\tilde\alpha)\right|\\ &\le 4C\|\eta-\eta_0\|_{\mathcal{N}}\|\theta-\theta_0\|_{\mathcal{H}}\rho_{\mathcal{H}}(\alpha)\\ &= O(\|\eta-\eta_0\|_{\mathcal{N}}\|\theta-\theta_0\|_{\mathcal{H}})\end{aligned} \tag{C.5}$$

Hence,

$$\begin{aligned}\dot\psi_0(\theta_0)(\theta-\theta_0) - \partial_\theta L_0(\theta,\eta)(\alpha) = \partial_\theta^2 L_0(\theta_0,\eta_0)(\alpha_0-\alpha,\theta-\theta_0)\\ + O(\|\theta-\theta_0\|_{\mathcal{H}}^2) + O(\|\eta-\eta_0\|_{\mathcal{N}}^2) + O(\|\eta-\eta_0\|_{\mathcal{N}}\|\theta-\theta_0\|_{\mathcal{H}}).\end{aligned}$$

Combining the previous expression with (C.1) and (C.2), this implies that

$$\begin{aligned}\psi_0(\theta) - \psi_0(\theta_0) - P_0\dot\ell_\eta(\theta)(\alpha) = \partial_\theta^2 L_0(\theta_0,\eta_0)(\alpha_0-\alpha,\theta-\theta_0) + O\left(\|\theta-\theta_0\|_{\mathcal{H}}^2\right)\\ + O(\|\eta-\eta_0\|_{\mathcal{N}}\|\theta-\theta_0\|_{\mathcal{H}}) + O\left(\|\eta-\eta_0\|_{\mathcal{N}}^2\right).\end{aligned}$$

Here the $O\left(\|\theta-\theta_0\|_{\mathcal{H}}^2\right)$ term collects the second-order remainders from the functional and risk expansions, and is absent when the functional is linear and the risk is quadratic.

□

C.2.4 Proofs for pathwise differentiability and EIF

We begin with an informal derivation of Theorem 4.11. We then give a rigorous proof using the same differentiation argument more carefully.

Informal derivation of EIF for Theorem 4.11. For any bounded score $S \in L_0^2(P_0)$, define the smooth submodel through $P_0 \in \mathcal{P}$ as $dP_{0,\varepsilon} = (1+\varepsilon S(z))dP_0(z)$. The first-order optimality conditions of the M-estimand $\theta_{P_{0,\varepsilon}}$ imply that, for each direction $h \in \mathcal{H}$:

$$\partial_\theta L_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}}, \eta_{0,P_{0,\varepsilon}})(h) = 0. \tag{C.6}$$

Taking the derivative of both sides, we find that

$$\frac{d}{d\varepsilon}\partial_\theta L_{P_{0,\varepsilon}}(\theta_0,\eta_0)(h)\big|_{\varepsilon=0} + \frac{d}{d\varepsilon}\partial_\theta L_0(\theta_{P_{0,\varepsilon}},\eta_0)(h)\big|_{\varepsilon=0} + \frac{d}{d\varepsilon}\partial_\theta L_0(\theta_0,\eta_{P_{0,\varepsilon}})(h)\big|_{\varepsilon=0} = 0.$$

By calculations given in the rigorous version of this proof, the first term equals $\langle \dot{\ell}_{\eta_0}(\theta_0)(h), S\rangle$, the second term equals $\partial_\theta^2 L_0(\theta_0,\eta_0)\left(\frac{d}{d\varepsilon}\theta_{P_{0,\varepsilon}}\big|_{\varepsilon=0}, h\right)$, and the final term is zero by Neyman-orthogonality of the loss. Rearranging terms, we find that

$$\partial_\theta^2 L_0(\theta_0,\eta_0)\left(\frac{d}{d\varepsilon}\theta_{P_{0,\varepsilon}}\big|_{\varepsilon=0}, h\right) = -\langle \dot{\ell}_{\eta_0}(\theta_0)(h), S\rangle_{L^2(P_0)}. \tag{C.7}$$

Differentiating the functional and applying the chain rule, we have

$$\frac{d}{d\varepsilon}\psi_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}})\big|_{\varepsilon=0} = \dot{\psi}_0(\theta_0)\left(\frac{d}{d\varepsilon}\theta_{P_{0,\varepsilon}}\big|_{\varepsilon=0}\right) + \langle m(\cdot,\theta_0) - \Psi(P_0), S\rangle_{L^2(P_0)}. \tag{C.8}$$

By the Riesz representation property in (4.2), we have

$$\dot{\psi}_0(\theta_0)\left(\frac{d}{d\varepsilon}\theta_{P_{0,\varepsilon}}\big|_{\varepsilon=0}\right) = \partial_\theta^2 L_0(\theta_0,\eta_0)\left(\frac{d}{d\varepsilon}\theta_{P_{0,\varepsilon}}\big|_{\varepsilon=0}, \alpha_0\right).$$

Applying (C.7) and plugging the above into (C.8), we conclude

$$\frac{d}{d\varepsilon}\psi_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}})\big|_{\varepsilon=0} = \langle -\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0) + m(\cdot,\theta_0) - \Psi(P_0), S\rangle_{L^2(P_0)}.$$

Since the above holds for any bounded score S, we can show that $\chi_0 := -\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0) + m(\cdot,\theta_0) - \Psi(P_0)$ is the efficient influence function of Ψ. □

Proof of Theorem 4.11. By A1, the M-estimand θ_P is well defined for each $P \in \mathcal{P}$. For any bounded score $S \in L_0^2(P_0)$, define the smooth submodel through P_0 as $dP_{0,\varepsilon} = (1 + \varepsilon S(z))\, dP_0(z)$ with $\varepsilon \in (-\delta, \delta)$ for $\delta > 0$ sufficiently small. Since $\mathcal{P}$ is convex by assumption, $P_{0,\varepsilon} \in \mathcal{P}$. Condition A3i and (4.1) imply that, for ε in a neighborhood of zero, $\theta_{P_{0,\varepsilon}}$ satisfies the following first-order optimality condition:

$$\partial_\theta L_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}}, \eta_{0,P_{0,\varepsilon}})(\alpha_0) = 0. \tag{C.9}$$

Applying the above and then adding and subtracting terms,

$$\begin{aligned} 0 = \partial_\theta L_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}}, \eta_{P_{0,\varepsilon}})(\alpha_0) - \partial_\theta L_0(\theta_0,\eta_0)(\alpha_0) &= \left\{\partial_\theta L_0(\theta_0, \eta_{P_{0,\varepsilon}})(\alpha_0) - \partial_\theta L_0(\theta_0,\eta_0)(\alpha_0)\right\} \\ &+ \left\{\partial_\theta L_0(\theta_{P_{0,\varepsilon}}, \eta_{P_{0,\varepsilon}})(\alpha_0) - \partial_\theta L_0(\theta_0, \eta_{P_{0,\varepsilon}})(\alpha_0)\right\} \\ &+ \left\{\partial_\theta L_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}}, \eta_{P_{0,\varepsilon}})(\alpha_0) - \partial_\theta L_0(\theta_{P_{0,\varepsilon}}, \eta_{P_{0,\varepsilon}})(\alpha_0)\right\}. \end{aligned} \tag{C.10}$$

In what follows, we will show that the three differences on the right-hand side are, respectively, $o(\varepsilon)$, $\partial_\theta^2 L(\theta_0,\eta_0)(\theta_{P_{0,\varepsilon}} - \theta_0, \alpha_0) + o(\varepsilon)$, and $\varepsilon\langle \dot{\ell}_{\eta_0}(\theta_0)(\alpha_0), S\rangle_{L^2(P_0)} + o(\varepsilon)$. Consequently, (C.10) implies that $\langle \dot{\ell}_{\eta_0}(\theta_0)(\alpha_0), S\rangle_{L^2(P_0)} = -\partial_\theta^2 L(\theta_0,\eta_0)(\varepsilon^{-1}(\theta_{P_{0,\varepsilon}} - \theta_0), \alpha_0) + o(1)$.

To this end, by A4 and Lemma C.3, the first difference on the right-hand side of (C.10) satisfies

$$\partial_\theta L_0(\theta_0, \eta_{P_{0,\varepsilon}})(\alpha_0) - \partial_\theta L_0(\theta_0,\eta_0)(\alpha_0) = \partial_\eta \partial_\theta L_0(\theta_0,\eta_0)(\eta_{P_{0,\varepsilon}} - \eta_0, \alpha_0) + O\left(\|\eta_{P_{0,\varepsilon}} - \eta_0\|_{\mathcal{N}}^2\right).$$

By A5, we have $\partial_\eta \partial_\theta L_0(\theta_0,\eta_0)(\eta_{P_{0,\varepsilon}} - \eta_0, \alpha_0) = 0$, and by the Lipschitz Hellinger continuity of $P \mapsto \eta_P$ in Condition A7, as $\varepsilon \to 0$, $\|\eta_{P_{0,\varepsilon}} - \eta_0\|_{\mathcal{N}}^2 = O(\varepsilon^2) = o(\varepsilon)$. Hence,

$$\partial_\theta L_0(\theta_0, \eta_{P_{0,\varepsilon}})(\alpha_0) - \partial_\theta L_0(\theta_0,\eta_0)(\alpha_0) = o(\varepsilon). \tag{C.11}$$

The second difference on the right-hand side of (C.10) satisfies, by A3 and Lemma C.2,

$$\begin{aligned}\partial_\theta L_0(\theta_{P_{0,\varepsilon}},\eta_{P_{0,\varepsilon}})(\alpha_0)-\partial_\theta L_0(\theta_0,\eta_{P_{0,\varepsilon}})(\alpha_0) &= \partial_\theta^2 L_0(\theta_{P_{0,\varepsilon}},\eta_{P_{0,\varepsilon}})(\theta_{P_{0,\varepsilon}}-\theta_0,\alpha_0)+O\left(\|\theta_{P_{0,\varepsilon}}-\theta_0\|_{\mathcal{H}}^2\right)\\ &= \partial_\theta^2 L_0(\theta_0,\eta_0)(\theta_{P_{0,\varepsilon}}-\theta_0,\alpha_0)+O\left(\|\theta_{P_{0,\varepsilon}}-\theta_0\|_{\mathcal{H}}^2\right)\\ &\quad+\left\{\partial_\theta^2 L_0(\theta_{P_{0,\varepsilon}},\eta_{P_{0,\varepsilon}})(\theta_{P_{0,\varepsilon}}-\theta_0,\alpha_0)-\partial_\theta^2 L_0(\theta_0,\eta_0)(\theta_{P_{0,\varepsilon}}-\theta_0,\alpha_0)\right\}.\end{aligned}$$

By the Lipschitz Hellinger continuity of $P\mapsto\theta_P$ and $P\mapsto\eta_P$ in Condition A7, we have $\|\theta_{P_{0,\varepsilon}}-\theta_0\|_{\mathcal{H}}=O(\varepsilon)$ and $\|\eta_{P_{0,\varepsilon}}-\eta_0\|_{\mathcal{N}}=O(\varepsilon)$. Moreover, if $\theta_{P_{0,\varepsilon}}=\theta_0$ or $\rho_{\mathcal{H}}(\alpha_0)=0$, then the same bound is immediate. Otherwise, combining this with the Lipschitz continuity of the second derivative in Condition A3iii, and writing

$$\tilde\theta_\varepsilon:=\frac{\theta_{P_{0,\varepsilon}}-\theta_0}{2\|\theta_{P_{0,\varepsilon}}-\theta_0\|_{\mathcal{H}}},\qquad \tilde\alpha_0:=\frac{\alpha_0}{2\rho_{\mathcal{H}}(\alpha_0)},$$

and

$$D_\varepsilon:=\partial_\theta^2 L_0(\theta_{P_{0,\varepsilon}},\eta_{P_{0,\varepsilon}})(\tilde\theta_\varepsilon,\tilde\alpha_0)-\partial_\theta^2 L_0(\theta_0,\eta_0)(\tilde\theta_\varepsilon,\tilde\alpha_0),$$

we obtain

$$\begin{aligned}\left|\partial_\theta^2 L_0(\theta_{P_{0,\varepsilon}},\eta_{P_{0,\varepsilon}})(\theta_{P_{0,\varepsilon}}-\theta_0,\alpha_0)-\partial_\theta^2 L_0(\theta_0,\eta_0)(\theta_{P_{0,\varepsilon}}-\theta_0,\alpha_0)\right| &= 4\|\theta_{P_{0,\varepsilon}}-\theta_0\|_{\mathcal{H}}\rho_{\mathcal{H}}(\alpha_0)\,|D_\varepsilon|\\ &\lesssim \rho_{\mathcal{H}}(\alpha_0)\|\theta_{P_{0,\varepsilon}}-\theta_0\|_{\mathcal{H}}\left\{\|\theta_{P_{0,\varepsilon}}-\theta_0\|_{\mathcal{H}}+\|\eta_{P_{0,\varepsilon}}-\eta_0\|_{\mathcal{N}}\right\}\\ &= O(\varepsilon^2)=o(\varepsilon).\end{aligned}$$

Thus,

$$\partial_\theta L_0(\theta_{P_{0,\varepsilon}},\eta_{P_{0,\varepsilon}})(\alpha_0)-\partial_\theta L_0(\theta_0,\eta_{P_{0,\varepsilon}})(\alpha_0)=\partial_\theta^2 L_0(\theta_0,\eta_0)(\theta_{P_{0,\varepsilon}}-\theta_0,\alpha_0)+o(\varepsilon). \tag{C.12}$$

The third difference on the right-hand side of (C.10) satisfies, by Condition A3i,

$$\begin{aligned}\partial_\theta L_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}},\eta_{P_{0,\varepsilon}})(\alpha_0)-\partial_\theta L_0(\theta_{P_{0,\varepsilon}},\eta_{P_{0,\varepsilon}})(\alpha_0) &= (P_{0,\varepsilon}-P_0)\dot\ell_{\eta_{P_{0,\varepsilon}}}(\theta_{P_{0,\varepsilon}})(\alpha_0)\\ &= \varepsilon\int \dot\ell_{\eta_{P_{0,\varepsilon}}}(\theta_{P_{0,\varepsilon}})(\alpha_0)(z)S(z)\,P_0(dz)\\ &= \varepsilon\int \dot\ell_{\eta_0}(\theta_0)(\alpha_0)(z)S(z)\,P_0(dz)+\varepsilon\left\langle \dot\ell_{\eta_{P_{0,\varepsilon}}}(\theta_{P_{0,\varepsilon}})(\alpha_0)-\dot\ell_{\eta_0}(\theta_0)(\alpha_0),S\right\rangle_{L^2(P_0}\end{aligned}$$

By the Hellinger continuity of $P\mapsto\eta_P$ and $P\mapsto\theta_P$, the continuity of the map $(\eta,\theta)\mapsto\dot\ell_\eta(\cdot,\theta)$ in Condition A7, and the continuity of the inner product, it follows that, as $\varepsilon\to0$,

$$\left\langle \dot\ell_{\eta_{P_{0,\varepsilon}}}(\theta_{P_{0,\varepsilon}})(\alpha_0)-\dot\ell_{\eta_0}(\theta_0)(\alpha_0),S\right\rangle_{L^2(P_0)}\le\|\dot\ell_{\eta_{P_{0,\varepsilon}}}(\theta_{P_{0,\varepsilon}})(\alpha_0)-\dot\ell_{\eta_0}(\theta_0)(\alpha_0)\|_{L^2(P_0)}\|S\|_{L^2(P_0)}=o(1).$$

Hence,

$$\partial_\theta L_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}},\eta_{P_{0,\varepsilon}})(\alpha_0)-\partial_\theta L_0(\theta_{P_{0,\varepsilon}},\eta_{P_{0,\varepsilon}})(\alpha_0)=\varepsilon\langle\dot\ell_{\eta_0}(\theta_0)(\alpha_0),S\rangle_{L^2(P_0)}+o(\varepsilon). \tag{C.13}$$

Finally, plugging the expressions in (C.11), (C.12), and (C.13) into (C.10), we obtain

$$0=\partial_\theta L_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}},\eta_{P_{0,\varepsilon}})(\alpha_0)-\partial_\theta L_0(\theta_0,\eta_0)(\alpha_0)=\partial_\theta^2 L_0(\theta_0,\eta_0)(\theta_{P_{0,\varepsilon}}-\theta_0,\alpha_0)+\varepsilon\langle\dot\ell_{\eta_0}(\theta_0)(\alpha_0),S\rangle_{L^2(P_0)}+o(\varepsilon).$$

Rearranging terms and dividing both sides by ε yields

$$\partial_\theta^2 L_0(\theta_0,\eta_0)\left(\frac{\theta_{P_{0,\varepsilon}}-\theta_0}{\varepsilon},\alpha_0\right)=\langle-\dot\ell_{\eta_0}(\theta_0)(\alpha_0),S\rangle_{L^2(P_0)}+o(1). \tag{C.14}$$

We are now in a position to compute the pathwise derivative of Ψ. By A2, we have that

$$\begin{aligned}\psi_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}})-\psi_0(\theta_0) &= \psi_0(\theta_{P_{0,\varepsilon}})-\psi_0(\theta_0)+\psi_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}})-\psi_0(\theta_{P_{0,\varepsilon}})\\ &= \varepsilon\dot\psi_0(\theta_0)\left(\varepsilon^{-1}\{\theta_{P_{0,\varepsilon}}-\theta_0\}\right)+o\left(\|\theta_{P_{0,\varepsilon}}-\theta_0\|_{\mathcal{H}}\right)+\varepsilon\langle m(\cdot,\theta_{P_{0,\varepsilon}}),S\rangle_{L^2(P_0)},\end{aligned}$$

where we use that $\psi_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}}) - \psi_0(\theta_{P_{0,\varepsilon}}) = \varepsilon\langle m(\cdot,\theta_{P_{0,\varepsilon}}), S\rangle_{L^2(P_0)}$ by the definition of ψ_P and the choice of submodel. Moreover, by the Cauchy-Schwarz inequality and A7, we have $\|\theta_{P_{0,\varepsilon}} - \theta_0\|_{\mathcal{H}} = O(\varepsilon)$ and $\langle m(\cdot,\theta_{P_{0,\varepsilon}}) - m(\cdot,\theta_0), S\rangle_{L^2(P_0)} \leq \|m(\cdot,\theta_{P_{0,\varepsilon}}) - m(\cdot,\theta_0)\|_{L^2(P_0)}\|S\|_{L^2(P_0)} = o(1)$ as $\varepsilon \to 0$. Hence,

$$\psi_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}}) - \psi_0(\theta_0) = \varepsilon\dot{\psi}_0(\theta_0)\left(\varepsilon^{-1}\{\theta_{P_{0,\varepsilon}} - \theta_0\}\right) + \varepsilon\langle m(\cdot,\theta_0), S\rangle_{L^2(P_0)} + o(\varepsilon).$$

Conditions A3ii and A6 imply that $\partial_\theta^2 L_0(\theta_0,\eta_0)(\cdot,\cdot)$ defines an inner product on $\overline{\mathcal{H}}$, and Condition A2 combined with A6 ensures that the functional derivative $\dot{\psi}_0(\theta_0)$ is continuous with respect to this inner product. Hence, the Riesz representer α_0 exists, and it satisfies the Riesz representation property:

$$\dot{\psi}_0(\theta_0)(h') = \partial_\theta^2 L_0(\theta_0,\eta_0)(h',\alpha_0).$$

Taking $h' = \varepsilon^{-1}\{\theta_{P_{0,\varepsilon}} - \theta_0\}$ in the above display, we find that

$$\psi_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}}) - \psi_0(\theta_0) = \varepsilon\partial_\theta^2 L_0(\theta_0,\eta_0)\left(\varepsilon^{-1}\{\theta_{P_{0,\varepsilon}} - \theta_0\},\alpha_0\right) + \varepsilon\langle m(\cdot,\theta_0), S\rangle_{L^2(P_0)} + o(\varepsilon).$$

Substituting (C.14) into the above expression, this implies that

$$\psi_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}}) - \psi_0(\theta_0) = \varepsilon\langle -\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0), S\rangle_{L^2(P_0)} + \varepsilon\langle m(\cdot,\theta_0), S\rangle_{L^2(P_0)} + o(\varepsilon).$$

Hence,

$$\begin{aligned}\lim_{\varepsilon\to 0}\frac{\psi_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}}) - \psi_0(\theta_0)}{\varepsilon}\bigg|_{\varepsilon=0} &= \langle -\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0), S\rangle_{L^2(P_0)} + \langle m(\cdot,\theta_0) - \Psi(P_0), S\rangle_{L^2(P_0)} \\ &= \langle -\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0) + m(\cdot,\theta_0) - \Psi(P_0), S\rangle_{L^2(P_0)}.\end{aligned}$$

The above holds for any bounded score S. Since the linear space of uniformly bounded scores is dense in $L_0^2(P_0)$ and Ψ is Hellinger Lipschitz continuous by A7, the above also holds, by continuity of the inner product, for any score $S \in L_0^2(P_0)$ [Lemma 2 of Luedtke and Chung, 2024]. We conclude that $\chi_0 = -\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0) + m(\cdot,\theta_0) - \Psi(P_0)$ is a gradient at P_0 for the pathwise derivative of Ψ. Since the model is nonparametric, χ_0 is the efficient influence function.

□

C.2.5 Pathwise derivative of M-estimand

In the following theorem, let $\dot{\ell}^*_{\eta_0}(\theta_0) : L_0^2(P_0) \to \overline{\mathcal{H}}$ be the adjoint of $\dot{\ell}_{\eta_0}(\theta_0) : \overline{\mathcal{H}} \to L_0^2(P_0)$, where $\overline{\mathcal{H}}$ is the closure of $\mathcal{H}$ equipped with the inner product $\partial_\theta^2 L_0(\theta_0,\eta_0)$.

Theorem C.5 (Pathwise derivative of M-estimand)**.** *Assume conditions A1-A8. Suppose that $P \mapsto \theta_P$ is pathwise differentiable at P_0 in a Hilbert-valued sense [Luedtke and Chung, 2024] with derivative $\dot{\theta}_0 : L_0^2(P_0) \to \overline{\mathcal{H}}$. Then, $\dot{\theta}_0 = -\dot{\ell}^*_{\eta_0}(\theta_0)$ and $\partial_\theta^2 L_0(\theta_0,\eta_0)(\dot{\theta}_0(S),h) = \langle -\dot{\ell}_{\eta_0}(\theta_0)(h), S\rangle_{L^2(P_0)}$.*

Proof. Fix a bounded score $S \in L_0^2(P_0)$ and define the submodel $dP_{0,\varepsilon} = (1+\varepsilon S)\,dP_0$. For any $h \in \mathcal{H}$, the first-order condition implies

$$\partial_\theta L_{P_{0,\varepsilon}}(\theta_{P_{0,\varepsilon}}, \eta_{P_{0,\varepsilon}})(h) = 0.$$

Repeating the decomposition in (C.10), but with h in place of α_0, and using the same bounds as in the proof of Theorem 4.11, we obtain

$$\partial_\theta^2 L_0(\theta_0, \eta_0)\left(\frac{\theta_{P_{0,\varepsilon}} - \theta_0}{\varepsilon}, h\right) = \langle -\dot{\ell}_{\eta_0}(\theta_0)(h), S\rangle_{L^2(P_0)} + o(1).$$

By pathwise differentiability of $P \mapsto \theta_P$, we have $\varepsilon^{-1}(\theta_{P_{0,\varepsilon}} - \theta_0) \to \dot{\theta}_0(S)$ in $\overline{\mathcal{H}}$ as $\varepsilon \to 0$. Passing to the limit yields

$$\partial_\theta^2 L_0(\theta_0, \eta_0)\left(\dot{\theta}_0(S), h\right) = \langle -\dot{\ell}_{\eta_0}(\theta_0)(h), S\rangle_{L^2(P_0)}$$

for all $h \in \mathcal{H}$. Since both sides are continuous in h with respect to the Hessian norm, the identity extends to all $h \in \overline{\mathcal{H}}$. By the definition of the adjoint,

$$\partial_\theta^2 L_0(\theta_0, \eta_0)\left(\dot{\ell}^*_{\eta_0}(\theta_0)(S), h\right) = \langle \dot{\ell}_{\eta_0}(\theta_0)(h), S\rangle_{L^2(P_0)}.$$

Therefore,

$$\partial_\theta^2 L_0(\theta_0, \eta_0)\left(\dot{\theta}_0(S) + \dot{\ell}^*_{\eta_0}(\theta_0)(S), h\right) = 0$$

for all $h \in \overline{\mathcal{H}}$. Taking $h = \dot{\theta}_0(S) + \dot{\ell}^*_{\eta_0}(\theta_0)(S)$ and using Condition A6, we conclude that

$$\dot{\theta}_0(S) = -\dot{\ell}^*_{\eta_0}(\theta_0)(S).$$

Since bounded scores are dense in $L_0^2(P_0)$ and both sides define continuous linear maps in S, the identity extends to all $S \in L_0^2(P_0)$. □

Assuming that $P \mapsto \theta_P$ is pathwise differentiable at P_0, we can derive the EIF of a smooth functional ψ of θ_P using the chain rule [Luedtke and Chung, 2024]. Specifically, the pathwise derivative of $P \mapsto \psi(\theta_P)$ is $\dot{\psi}(\theta_0) \circ \dot{\theta}_0 = -\dot{\psi}(\theta_0) \circ \dot{\ell}^*_{\eta_0}(\theta_0)$. Suppose that $\dot{\psi}(\theta_0)(h) = \partial_\theta^2 L_0(\theta_0, \eta_0)(\alpha_0, h)$ for a Riesz representer α_0. Then,

$$\begin{aligned}(\dot{\psi}(\theta_0) \circ \dot{\theta}_0)(S) &= \partial_\theta^2 L_0(\theta_0, \eta_0)\left(\alpha_0, -\dot{\ell}^*_{\eta_0}(\theta_0)(S)\right) \\ &= \langle -\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0), S\rangle_{L^2(P_0)}\end{aligned}$$

where $-\dot{\ell}_{\eta_0}(\theta_0)(\alpha_0)$ is the efficient influence function of $P \mapsto \psi(\theta_P)$ at P_0.

C.3 Extensions and refinements of the von Mises expansion

C.3.1 Constrained Lagrangian von Mises expansion

We record the analogue of Theorem 4.9 for equality-constrained M-estimands. Let $\mathcal{C}$ be a Hilbert space and define

$$C_0(\theta, \eta) := P_0 c_\eta(\cdot, \theta) \in \mathcal{C}, \qquad \Lambda_0(\theta, \eta, \lambda) := L_0(\theta, \eta) + \langle \lambda, C_0(\theta, \eta)\rangle_{\mathcal{C}}.$$

Here $c_\eta(z, \theta)$ is $\mathcal{C}$-valued and $\lambda \in \mathcal{C}$ is a Lagrange multiplier. A $\mathcal{C}$-valued random element V is Bochner integrable under P_0 if it is strongly measurable and $P_0\|V\|_{\mathcal{C}} < \infty$; in that case $P_0 V$ denotes its $\mathcal{C}$-valued expectation. The constrained truth is

$$\theta_0 = \operatorname*{argmin}_{\theta \in \mathcal{H}} L_0(\theta, \eta_0) \qquad \text{subject to } C_0(\theta, \eta_0) = 0.$$

Fix a multiplier λ_0 satisfying the constrained first-order conditions below, and define

$$A_0 h := \partial_\theta C_0(\theta_0, \eta_0)(h), \qquad \mathcal{I}_0(h_1, h_2) := \partial_\theta^2 \Lambda_0(\theta_0, \eta_0, \lambda_0)(h_1, h_2).$$

If $\dot{c}_\eta(\theta)(h)$ denotes the pointwise derivative of $c_\eta(\cdot, \theta)$ in direction h, define the Lagrangian score

$$\dot{q}_{\eta,\lambda}(\theta)(h) := \dot{\ell}_\eta(\theta)(h) + \langle \lambda, \dot{c}_\eta(\theta)(h) \rangle_{\mathcal{C}}.$$

Then

$$P_0 \dot{q}_{\eta,\lambda_0}(\theta)(h) = \partial_\theta \Lambda_0(\theta, \eta, \lambda_0)(h).$$

The following conditions are the constrained counterparts of the assumptions used in Theorem 4.9.

C1) *(Constrained first-order conditions).* The constraint is feasible at (θ_0, η_0), and there exists $\lambda_0 \in \mathcal{C}$ such that

$$C_0(\theta_0, \eta_0) = 0, \qquad \partial_\theta \Lambda_0(\theta_0, \eta_0, \lambda_0)(h) = 0 \quad \text{for all } h \in \mathcal{H}.$$

C2) *(Constraint and score representation).*

i) For each $\theta \in \mathcal{H}_\mathcal{P}$ and $\eta \in \mathcal{N}_\mathcal{P}$, the $\mathcal{C}$-valued map $z \mapsto c_\eta(z, \theta)$ is Bochner integrable under P_0. Moreover, $\dot{\ell}_\eta(\theta)(h)$ is P_0-integrable and

$$P_0 \dot{\ell}_\eta(\theta)(h) = \partial_\theta L_0(\theta, \eta)(h) \qquad \text{for all } h \in \mathcal{H}.$$

ii) The map $\theta \mapsto C_0(\theta, \eta_0)$ is Fréchet differentiable at θ_0, with bounded derivative $A_0 : \mathcal{H} \to \mathcal{C}$.

iii) For each $\theta \in \mathcal{H}_\mathcal{P}$, $\eta \in \mathcal{N}_\mathcal{P}$, and $h \in \mathcal{H}$, the pointwise derivative $\dot{c}_\eta(\theta)(h)$ is Bochner integrable under P_0 and satisfies

$$P_0 \dot{c}_\eta(\theta)(h) = \partial_\theta C_0(\theta, \eta)(h).$$

C3) *(Hölder smoothness of the Lagrangian in θ).*

i) For each $\eta \in \mathcal{N}_\mathcal{P}$, the map $\partial_\theta \Lambda_0(\cdot, \eta, \lambda_0) : (\mathcal{H}, \|\cdot\|_\mathcal{H}) \to (\mathcal{H}, \rho_\mathcal{H})^*$ is Fréchet differentiable over $\mathcal{H}_\mathcal{P}$.

ii) The second derivative $\partial_\theta^2 \Lambda_0(\theta, \eta, \lambda_0)$ is symmetric.

iii) There exists $C < \infty$ such that, for all $\theta, \theta' \in \mathcal{H}_\mathcal{P}$, $\eta, \eta' \in \mathcal{N}_\mathcal{P}$, and $h_1, h_2 \in \mathcal{H}$ with $\|h_1\|_\mathcal{H} + \rho_\mathcal{H}(h_2) \le 1$,

$$\left|\partial_\theta^2 \Lambda_0(\theta', \eta', \lambda_0)(h_1, h_2) - \partial_\theta^2 \Lambda_0(\theta, \eta, \lambda_0)(h_1, h_2)\right| \le C \left\{\|\theta' - \theta\|_\mathcal{H} + \|\eta' - \eta\|_\mathcal{N}\right\}.$$

C4) *(Hölder smoothness of the Lagrangian in the nuisance).*

i) The map $\eta \mapsto \partial_\theta \Lambda_0(\theta_0, \eta, \lambda_0)$, defined on $\mathcal{N}_\mathcal{P}$, is Fréchet differentiable in η, with derivative

$$\partial_\eta \partial_\theta \Lambda_0(\theta_0, \eta, \lambda_0) : \widetilde{\mathcal{N}} \to (\mathcal{H}, \rho_\mathcal{H})^*.$$

ii) There exists $C < \infty$ such that, for all $\eta, \eta' \in \mathcal{N}_\mathcal{P}$, $g \in \widetilde{\mathcal{N}}$, and $h \in \mathcal{H}$ satisfying $\|g\|_\mathcal{N} + \rho_\mathcal{H}(h) \le 1$,

$$\left|\partial_\eta \partial_\theta \Lambda_0(\theta_0, \eta', \lambda_0)(g, h) - \partial_\eta \partial_\theta \Lambda_0(\theta_0, \eta, \lambda_0)(g, h)\right| \le C \|\eta' - \eta\|_\mathcal{N}.$$

C5) *(Lagrangian Neyman orthogonality).* For each $\eta \in \mathcal{N}_{\mathcal{P}}$ and $h \in \mathcal{H}$,

$$\partial_\eta \partial_\theta \Lambda_0(\theta_0, \eta_0, \lambda_0)(\eta - \eta_0, h) = 0.$$

C6) *(Constrained Riesz representation).* There exist $\alpha_0 \in \overline{\mathcal{H}}$ and $\gamma_0 \in \mathcal{C}$, with $\rho_{\mathcal{H}}(\alpha_0) < \infty$, such that $\mathcal{I}_0$ extends continuously to α_0 in its first argument and

$$\dot{\psi}_0(\theta_0)(h) = \mathcal{I}_0(\alpha_0, h) + \langle \gamma_0, A_0 h \rangle_{\mathcal{C}} \qquad \text{for all } h \in \mathcal{H}.$$

Condition C6 is the constrained analogue of (4.2). A standard sufficient condition is solvability of the linearized KKT system

$$\begin{pmatrix} \mathcal{I}_0 & A_0^* \\ A_0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \gamma_0 \end{pmatrix} = \begin{pmatrix} \dot{\psi}_0(\theta_0) \\ 0 \end{pmatrix},$$

where the first block equation is interpreted after testing against each $h \in \mathcal{H}$, as in the preceding display. Thus Condition C6 decomposes the target derivative into a component represented by the Lagrangian curvature and a component in the linearized constraint direction. The theorem does not require uniqueness of λ_0 or of (α_0, γ_0); it holds for any fixed choices satisfying Conditions C1 and C6.

Lemma C.6 (Quadratic expansions for the Lagrangian gradient)**.** *Suppose Conditions C3 and C4 hold. Then, uniformly over $\eta \in \mathcal{N}_{\mathcal{P}}$, $\theta \in \mathcal{H}_{\mathcal{P}}$, and $\alpha \in \mathcal{H}$ with $\rho_{\mathcal{H}}(\alpha) \le M$,*

$$\begin{aligned} \partial_\theta \Lambda_0(\theta, \eta, \lambda_0)(\alpha) = \partial_\theta \Lambda_0(\theta_0, \eta, \lambda_0)(\alpha) + \partial_\theta^2 \Lambda_0(\theta_0, \eta, \lambda_0)(\theta - \theta_0, \alpha) \\ + O\big(\|\theta - \theta_0\|_{\mathcal{H}}^2\big), \end{aligned}$$

where this remainder is absent when $\Lambda_0(\cdot, \eta, \lambda_0)$ is quadratic at θ_0, uniformly over $\eta \in \mathcal{N}_{\mathcal{P}}$. Also, uniformly over $\eta \in \mathcal{N}_{\mathcal{P}}$ and such α,

$$\begin{aligned} \partial_\theta \Lambda_0(\theta_0, \eta, \lambda_0)(\alpha) = \partial_\theta \Lambda_0(\theta_0, \eta_0, \lambda_0)(\alpha) + \partial_\eta \partial_\theta \Lambda_0(\theta_0, \eta_0, \lambda_0)(\eta - \eta_0, \alpha) \\ + O\big(\|\eta - \eta_0\|_{\mathcal{N}}^2\big). \end{aligned}$$

Proof. Let $\delta := \theta - \theta_0$. Taylor's theorem and Condition C3 apply along the segment $\theta_0 + t\delta \in \mathcal{H}_{\mathcal{P}}$. Using symmetry of the Lagrangian Hessian, they give

$$\partial_\theta \Lambda_0(\theta, \eta, \lambda_0)(\alpha) = \partial_\theta \Lambda_0(\theta_0, \eta, \lambda_0)(\alpha) + \partial_\theta^2 \Lambda_0(\theta_0, \eta, \lambda_0)(\delta, \alpha) + R_\theta,$$

where

$$R_\theta = \int_0^1 \left\{ \partial_\theta^2 \Lambda_0(\theta_0 + t\delta, \eta, \lambda_0)(\delta, \alpha) - \partial_\theta^2 \Lambda_0(\theta_0, \eta, \lambda_0)(\delta, \alpha) \right\} dt.$$

If $\|\delta\|_{\mathcal{H}} = 0$ or $\rho_{\mathcal{H}}(\alpha) = 0$, the bound is immediate. Otherwise, applying Condition C3 to $\delta/(2\|\delta\|_{\mathcal{H}})$ and $\alpha/(2\rho_{\mathcal{H}}(\alpha))$ yields

$$|R_\theta| \le C \rho_{\mathcal{H}}(\alpha) \|\delta\|_{\mathcal{H}}^2.$$

This proves the first expansion, and the remainder is zero when $\Lambda_0(\cdot, \eta, \lambda_0)$ is quadratic at θ_0.

Let $\Delta := \eta - \eta_0$. Taylor's theorem and Condition C4 similarly give

$$\partial_\theta \Lambda_0(\theta_0, \eta, \lambda_0)(\alpha) = \partial_\theta \Lambda_0(\theta_0, \eta_0, \lambda_0)(\alpha) + \partial_\eta \partial_\theta \Lambda_0(\theta_0, \eta_0, \lambda_0)(\Delta, \alpha) + R_\eta,$$

where

$$R_\eta = \int_0^1 \left\{ \partial_\eta \partial_\theta \Lambda_0(\theta_0, \eta_0 + t\Delta, \lambda_0)(\Delta, \alpha) - \partial_\eta \partial_\theta \Lambda_0(\theta_0, \eta_0, \lambda_0)(\Delta, \alpha) \right\} dt.$$

If $\|\Delta\|_{\mathcal{N}} = 0$ or $\rho_{\mathcal{H}}(\alpha) = 0$, the bound is immediate. Otherwise, Condition C4 applied to $\Delta/(2\|\Delta\|_{\mathcal{N}})$ and $\alpha/(2\rho_{\mathcal{H}}(\alpha))$ yields

$$|R_\eta| \le C\rho_{\mathcal{H}}(\alpha)\|\Delta\|_{\mathcal{N}}^2.$$

The condition $\rho_{\mathcal{H}}(\alpha) \le M$ absorbs the seminorm factors into the constants in the two $O(\cdot)$ terms. □

Theorem C.7 (Constrained Lagrangian von Mises expansion)**.** *Suppose Condition A2 and Conditions C1–C6 hold. Then, for each $\eta \in \mathcal{N}_{\mathcal{P}}$, $\theta \in \mathcal{H}_{\mathcal{P}}$, and $\alpha \in \mathcal{H}$ with $\rho_{\mathcal{H}}(\alpha) < M$,*

$$\begin{aligned}\psi_0(\theta) - \psi_0(\theta_0) = P_0\dot{q}_{\eta,\lambda_0}(\theta)(\alpha) + \mathcal{I}_0(\alpha_0 - \alpha, \theta - \theta_0) + \langle\gamma_0, A_0(\theta - \theta_0)\rangle_{\mathcal{C}} \\ + O\big(\|\theta - \theta_0\|_{\mathcal{H}}^2\big) + O(\|\eta - \eta_0\|_{\mathcal{N}}\|\theta - \theta_0\|_{\mathcal{H}}) + O\big(\|\eta - \eta_0\|_{\mathcal{N}}^2\big)\,,\end{aligned}$$

where the constants in the $O(\cdot)$ terms depend only on M, $\|\gamma_0\|_{\mathcal{C}}$, $\rho_{\mathcal{H}}(\alpha_0)$, and the constants in the preceding conditions. If ψ_0 is linear at θ_0 and $\Lambda_0(\cdot, \eta, \lambda_0)$ is quadratic at θ_0, uniformly over $\eta \in \mathcal{N}_{\mathcal{P}}$, then the Taylor remainders coming from the target and the Lagrangian are absent.

Proof. Let $\delta := \theta - \theta_0$. By Lemma C.1 and Condition A2,

$$\psi_0(\theta) - \psi_0(\theta_0) = \dot{\psi}_0(\theta_0)(\delta) + O\big(\|\delta\|_{\mathcal{H}}^2\big)\,,$$

with the remainder absent when ψ_0 is linear at θ_0. By Condition C6,

$$\dot{\psi}_0(\theta_0)(\delta) = \mathcal{I}_0(\alpha_0, \delta) + \langle\gamma_0, A_0\delta\rangle_{\mathcal{C}}.$$

It remains to expand the Lagrangian gradient in the direction α. Condition C2 and Lemma C.6 give

$$\begin{aligned}P_0\dot{q}_{\eta,\lambda_0}(\theta)(\alpha) &= \partial_\theta\Lambda_0(\theta, \eta, \lambda_0)(\alpha) \\ &= \partial_\theta\Lambda_0(\theta_0, \eta, \lambda_0)(\alpha) + \partial_\theta^2\Lambda_0(\theta_0, \eta, \lambda_0)(\delta, \alpha) + O\big(\|\delta\|_{\mathcal{H}}^2\big)\,,\end{aligned}$$

where the last remainder is absent when $\Lambda_0(\cdot, \eta, \lambda_0)$ is quadratic at θ_0, uniformly over $\eta \in \mathcal{N}_{\mathcal{P}}$. Expanding the first term in η, again using Lemma C.6, yields

$$\begin{aligned}\partial_\theta\Lambda_0(\theta_0, \eta, \lambda_0)(\alpha) = \partial_\theta\Lambda_0(\theta_0, \eta_0, \lambda_0)(\alpha) + \partial_\eta\partial_\theta\Lambda_0(\theta_0, \eta_0, \lambda_0)(\eta - \eta_0, \alpha) \\ + O\big(\|\eta - \eta_0\|_{\mathcal{N}}^2\big)\,.\end{aligned}$$

The two leading terms in this display are zero by Conditions C1 and C5. Therefore,

$$P_0\dot{q}_{\eta,\lambda_0}(\theta)(\alpha) = \partial_\theta^2\Lambda_0(\theta_0, \eta, \lambda_0)(\delta, \alpha) + O\big(\|\delta\|_{\mathcal{H}}^2\big) + O\big(\|\eta - \eta_0\|_{\mathcal{N}}^2\big)\,.$$

Condition C3, applied as in (C.5), gives

$$\partial_\theta^2\Lambda_0(\theta_0, \eta, \lambda_0)(\delta, \alpha) = \mathcal{I}_0(\delta, \alpha) + O(\|\eta - \eta_0\|_{\mathcal{N}}\|\delta\|_{\mathcal{H}})\,.$$

Since $\mathcal{I}_0$ is symmetric,

$$P_0\dot{q}_{\eta,\lambda_0}(\theta)(\alpha) = \mathcal{I}_0(\alpha, \delta) + O\big(\|\delta\|_{\mathcal{H}}^2\big) + O(\|\eta - \eta_0\|_{\mathcal{N}}\|\delta\|_{\mathcal{H}}) + O\big(\|\eta - \eta_0\|_{\mathcal{N}}^2\big)\,.$$

Combining this expansion with the target expansion and the constrained Riesz representation gives the stated result. □

For feasible constrained candidates, the additional linearized-constraint term is a second-order remainder under the following local smoothness condition.

C7) *(Second-order constraint expansion).* There exists $C < \infty$ such that, for each $\theta \in \mathcal{H}_{\mathcal{P}}$,

$$\|C_0(\theta, \eta_0) - C_0(\theta_0, \eta_0) - A_0(\theta - \theta_0)\|_{\mathcal{C}} \le C\|\theta - \theta_0\|_{\mathcal{H}}^2.$$

Corollary C.8 (Feasible constrained expansion)**.** *Assume the conditions of Theorem C.7 and Condition C7. If, in addition, θ is feasible, that is, $C_0(\theta, \eta_0) = 0$, then*

$$\begin{aligned}\psi_0(\theta) - \psi_0(\theta_0) = P_0\dot{q}_{\eta,\lambda_0}(\theta)(\alpha) + \mathcal{I}_0(\alpha_0 - \alpha, \theta - \theta_0) \\ + O\big(\|\theta - \theta_0\|_{\mathcal{H}}^2\big) + O(\|\eta - \eta_0\|_{\mathcal{N}}\|\theta - \theta_0\|_{\mathcal{H}}) + O\big(\|\eta - \eta_0\|_{\mathcal{N}}^2\big)\,.\end{aligned}$$

If ψ_0 is linear in θ, $\Lambda_0(\cdot, \eta, \lambda_0)$ is quadratic in θ, uniformly over $\eta \in \mathcal{N}_{\mathcal{P}}$, and $C_0(\cdot, \eta_0)$ is affine, then the $O(\|\theta - \theta_0\|_{\mathcal{H}}^2)$ term is absent.

Proof. Let $\delta := \theta - \theta_0$. Since $C_0(\theta, \eta_0) = 0$ and $C_0(\theta_0, \eta_0) = 0$, Condition C7 gives

$$\|A_0\delta\|_{\mathcal{C}} \le C\|\delta\|_{\mathcal{H}}^2.$$

Therefore

$$|\langle \gamma_0, A_0\delta\rangle_{\mathcal{C}}| \le \|\gamma_0\|_{\mathcal{C}}\|A_0\delta\|_{\mathcal{C}} = O\big(\|\delta\|_{\mathcal{H}}^2\big)\,.$$

Substituting this bound into Theorem C.7 gives the displayed expansion. If ψ_0 is linear in θ, the Lagrangian is uniformly quadratic in θ, and the constraint map is affine, then the Taylor remainders in θ and the linearized-constraint remainder vanish, so the $O(\|\delta\|_{\mathcal{H}}^2)$ term is absent. □

When $\mathcal{C} = \{0\}$, the Lagrangian reduces to L_0, $A_0 = 0$, and $\gamma_0 = 0$. Thus Theorem C.7 reduces to Theorem 4.9.

C.3.2 Universal Neyman-orthogonality refinement

The following theorem records a refinement of Theorem 4.9. If the Hessian is itself Neyman-orthogonal in the nuisance direction, then the Lipschitz step in (C.5) improves from first order to second order in $\|\eta - \eta_0\|_{\mathcal{N}}$.

Theorem C.9 (Von Mises expansion under universal Neyman-orthogonality)**.** *Assume the conditions of Theorem 4.9 and the following additional conditions.*

(i) (Universal Hessian orthogonality). For all $h_1, h_2 \in \mathcal{H}$ and $\eta \in \mathcal{N}_{\mathcal{P}}$,

$$\partial_\eta\partial_\theta^2 L_0(\theta_0, \eta_0)(\eta - \eta_0, h_1, h_2) = 0.$$

(ii) (Fréchet differentiability of the Hessian in η). For each $\theta \in \mathcal{H}_{\mathcal{P}}$, the map $\eta \mapsto \partial_\theta^2 L_0(\theta, \eta)$, defined on $\mathcal{N}_{\mathcal{P}}$, is Fréchet differentiable in η, with derivative

$$\partial_\eta\partial_\theta^2 L_0(\theta, \eta) : \widetilde{\mathcal{N}} \times \mathcal{H} \times \mathcal{H} \to \mathbb{R}.$$

(iii) (Lipschitz continuity of the Hessian cross-derivative). There exists $C < \infty$ such that, for all $\eta, \eta' \in \mathcal{N}_{\mathcal{P}}$, $g \in \widetilde{\mathcal{N}}$, and $h_1, h_2 \in \mathcal{H}$ with $\|g\|_{\mathcal{N}} + \|h_1\|_{\mathcal{H}} + \rho_{\mathcal{H}}(h_2) \le 1$,

$$\begin{aligned}&\big|\partial_\eta\partial_\theta^2 L_0(\theta_0, \eta')(g, h_1, h_2) - \partial_\eta\partial_\theta^2 L_0(\theta_0, \eta)(g, h_1, h_2)\big| \\ &\quad \le C\|\eta' - \eta\|_{\mathcal{N}}.\end{aligned}$$

Then, for each $\eta \in \mathcal{N}_{\mathcal{P}}$, $\theta \in \mathcal{H}_{\mathcal{P}}$, and $\alpha \in \mathcal{H}$ satisfying $\rho_{\mathcal{H}}(\alpha) < M$,

$$\psi_0(\theta) - \psi_0(\theta_0) = P_0\dot{\ell}_\eta(\theta)(\alpha) + \partial_\theta^2 L_0(\theta_0, \eta_0)(\alpha_0 - \alpha, \theta - \theta_0) + O\big(\|\theta - \theta_0\|_{\mathcal{H}}^2\big) + O\big(\|\eta - \eta_0\|_{\mathcal{N}}^2\big),$$

where the constants in the $O(\cdot)$ terms depend only on M, the local radius of $\mathcal{H}_{\mathcal{P}}$, and the constants in the preceding conditions. If ψ_0 is linear at θ_0 and $L_0(\cdot, \eta_0)$ is quadratic at θ_0, then the $O(\|\theta - \theta_0\|_{\mathcal{H}}^2)$ term is absent.

Proof. Let $\delta := \theta - \theta_0$, $\Delta := \eta - \eta_0$, and

$$F(\eta) := \partial_\theta^2 L_0(\theta_0, \eta)(\delta, \alpha).$$

Since $\mathcal{N}_{\mathcal{P}}$ is convex, $\eta_0 + s\Delta \in \mathcal{N}_{\mathcal{P}}$ for $s \in [0, 1]$. By assumption (ii), Taylor's theorem with integral remainder gives

$$F(\eta) - F(\eta_0) = \int_0^1 \partial_\eta F(\eta_0 + s\Delta)[\Delta]\, ds,$$

where

$$\partial_\eta F(\eta)[\Delta] = \partial_\eta \partial_\theta^2 L_0(\theta_0, \eta)(\Delta, \delta, \alpha).$$

Universal Hessian orthogonality yields $\partial_\eta F(\eta_0)[\Delta] = 0$, and hence

$$F(\eta) - F(\eta_0) = \int_0^1 \{\partial_\eta F(\eta_0 + s\Delta) - \partial_\eta F(\eta_0)\}[\Delta]\, ds.$$

If $\|\Delta\|_{\mathcal{N}} = 0$, $\|\delta\|_{\mathcal{H}} = 0$, or $\rho_{\mathcal{H}}(\alpha) = 0$, the desired bound is immediate. Otherwise, trilinearity and assumption (iii), applied to $\Delta/(3\|\Delta\|_{\mathcal{N}})$, $\delta/(3\|\delta\|_{\mathcal{H}})$, and $\alpha/(3\rho_{\mathcal{H}}(\alpha))$, imply that, for each $s \in [0, 1]$,

$$\begin{aligned} &|\{\partial_\eta F(\eta_0 + s\Delta) - \partial_\eta F(\eta_0)\}[\Delta]| \\ &\quad = \big|\{\partial_\eta \partial_\theta^2 L_0(\theta_0, \eta_0 + s\Delta) - \partial_\eta \partial_\theta^2 L_0(\theta_0, \eta_0)\}(\Delta, \delta, \alpha)\big| \\ &\quad \le 27Cs\, \|\Delta\|_{\mathcal{N}}^2 \|\delta\|_{\mathcal{H}} \rho_{\mathcal{H}}(\alpha). \end{aligned}$$

The factors $\|\delta\|_{\mathcal{H}}$ and $\rho_{\mathcal{H}}(\alpha)$ are bounded on the local set and by the condition $\rho_{\mathcal{H}}(\alpha) < M$, respectively. Therefore,

$$|F(\eta) - F(\eta_0)| \le \int_0^1 O\big(s\|\Delta\|_{\mathcal{N}}^2\big)\, ds = O\big(\|\Delta\|_{\mathcal{N}}^2\big).$$

Thus

$$\big|\partial_\theta^2 L_0(\theta_0, \eta)(\delta, \alpha) - \partial_\theta^2 L_0(\theta_0, \eta_0)(\delta, \alpha)\big| = O\big(\|\eta - \eta_0\|_{\mathcal{N}}^2\big),$$

and substituting this bound into the proof of Theorem 4.9 at (C.5) removes the cross-product remainder. □

C.3.3 Orthogonalized moments from Hessian Riesz representers

The same Hessian-representer construction can be applied to moment equations. Suppose that φ_0 is identified by

$$P_0 U_{\theta_0}(\cdot; \varphi_0) = 0,$$

where θ_0 denotes an infinite-dimensional nuisance parameter and φ is finite dimensional. For a scalar component of the moment, and componentwise for vector-valued moments, define

$$\psi_0^{\varphi}(\theta) := P_0 U_\theta(\cdot; \varphi).$$

If $\theta \mapsto \psi_0^{\varphi}(\theta)$ satisfies the smoothness condition used in Theorem 4.9, let α_0^{φ} denote a Hessian Riesz representer of $\dot{\psi}_0^{\varphi}(\theta_0)$, so that

$$\dot{\psi}_0^{\varphi}(\theta_0)(h) = \partial_\theta^2 L_0(\theta_0, \eta_0)(\alpha_0^{\varphi}, h) \qquad \text{for all } h \in \mathcal{H}.$$

Define the orthogonalized moment

$$\widetilde{U}_{\eta,\theta,\alpha^{\varphi}}(Z;\varphi) := U_\theta(Z;\varphi) - \dot{\ell}_\eta(\theta)(\alpha^{\varphi})(Z).$$

At $(\eta_0, \theta_0, \alpha_0^{\varphi})$, the derivative of its population mean with respect to θ is zero. Indeed, for every direction $h \in \mathcal{H}$,

$$\begin{aligned}\partial_\theta\{P_0\widetilde{U}_{\eta_0,\theta,\alpha_0^{\varphi}}(\cdot;\varphi)\}(h)\Big|_{\theta=\theta_0} &= \dot{\psi}_0^{\varphi}(\theta_0)(h) - \partial_\theta^2 L_0(\theta_0,\eta_0)(h,\alpha_0^{\varphi}) \\ &= 0,\end{aligned}$$

where the last equality uses the Hessian Riesz representation and symmetry of $\partial_\theta^2 L_0(\theta_0, \eta_0)$. Hence the nuisance sensitivity of the estimating equation is removed to first order, and φ_0 can be estimated by solving

$$P_n\widetilde{U}_{\eta_n,\theta_n,\alpha_n^{\varphi}}(\cdot;\varphi) = 0,$$

under the usual regularity conditions for the resulting empirical moment [van der Laan and Dudoit, 2003, Chernozhukov et al., 2018a, Chen et al., 2026]. The same correction can also be applied to the first-order conditions of a population risk to construct Neyman-orthogonal losses, as in Foster and Syrgkanis [2023], van der Laan et al. [2024a].

C.4 Proofs for Section 4.4

Denote the EIF estimate corresponding to θ_n, η_n, and α_n by $\chi_n := m(\cdot, \theta_n) - P_n m(\cdot, \theta_n) - \dot{\ell}_{\eta_n}(\theta_n)(\alpha_n)$. For each $\eta \in \mathcal{N}_{\mathcal{P}}$, $\theta \in \mathcal{H}_{\mathcal{P}}$, and $\alpha \in \mathcal{H}$, define the remainder $\mathrm{Rem}_{P_0}(\eta, \theta, \alpha) := \psi_0(\theta) - \psi_0(\theta_0) - P_0\dot{\ell}_\eta(\theta)(\alpha)$, which by the von Mises expansion in Theorem 4.9 is second-order.

Lemma C.10 (Decomposition of bias of one-step estimator)**.** *Assume A1-A8 hold at* $P := P_0$*. Then,*

$$\widehat{\psi}_n^{\mathrm{dml}} - \Psi(P_0) = P_n\chi_0 + (P_n - P_0)(\chi_n - \chi_0) + Rem_{P_0}(\eta_n, \theta_n, \alpha_n).$$

Proof of Lemma C.10. Under conditions A1-A8, Ψ is a pathwise differentiable parameter at P_0 with EIF χ_0 provided in Theorem 4.11. By definition of $\mathrm{Rem}_{P_0}(\eta_n, \theta_n, \alpha_n)$, we have the following von Mises expansion of the bias:

$$P_n m(\cdot, \theta_n) - \Psi(P_0) = -P_0\chi_n + \mathrm{Rem}_{P_0}(\eta_n, \theta_n, \alpha_n),$$

noting that $\mathrm{Rem}_{P_0}(\eta_n, \theta_n, \alpha_n) = P_n m(\cdot, \theta_n) - \Psi(P_0) + P_0\chi_n$. Adding $P_n\chi_n$ to both sides, we find that

$$\begin{aligned}\widehat{\psi}_n^{\mathrm{dml}} - \Psi(P_0) &= P_n m(\cdot, \theta_n) + P_n\chi_n - \Psi(P_0) \\ &= (P_n - P_0)\chi_n + \mathrm{Rem}_{P_0}(\eta_n, \theta_n, \alpha_n) \\ &= (P_n - P_0)\chi_0 + (P_n - P_0)(\chi_n - \chi_0) + \mathrm{Rem}_{P_0}(\eta_n, \theta_n, \alpha_n).\end{aligned}$$

The result follows noting that $(P_n - P_0)\chi_0 = P_n\chi_0$, since $P_0\chi_0 = 0$. □

Proof of Theorem 4.13. By Lemma C.10, we have the bias expansion

$$\widehat{\psi}_n^{\mathrm{dml}} - \Psi(P_0) = P_n\chi_0 + (P_n - P_0)(\chi_n - \chi_0) + \mathrm{Rem}_{P_0}(\eta_n, \theta_n, \alpha_n).$$

By B5, we have $(P_n - P_0)(\chi_n - \chi_0) = o_p(n^{-1/2})$, so that

$$\widehat{\psi}_n^{\mathrm{dml}} - \Psi(P_0) = P_n\chi_0 + \mathrm{Rem}_{P_0}(\eta_n, \theta_n, \alpha_n) + o_p(n^{-1/2}).$$

Next, by B1, we can apply Theorem 4.9 to conclude that

$$\begin{aligned}\mathrm{Rem}_{P_0}(\eta_n, \theta_n, \alpha_n) = O_p\left(\partial_\theta^2 L_0(\theta_0, \eta_0)(\alpha_0 - \alpha_n, \theta_n - \theta_0)\right) \\ + O_p\left(\|\theta_n - \theta_0\|_{\mathcal{H}}^2\right) + O_p\left(\|\eta_n - \eta_0\|_{\mathcal{N}}^2\right) \\ + O_p\left(\|\eta_n - \eta_0\|_{\mathcal{N}}\|\theta_n - \theta_0\|_{\mathcal{H}}\right).\end{aligned}$$

By B2-B4, it holds that $\mathrm{Rem}_{P_0}(\eta_n, \theta_n, \alpha_n) = o_p(n^{-1/2})$. We conclude that

$$\widehat{\psi}_n^{\mathrm{dml}} - \Psi(P_0) = P_n\chi_0 + o_p(n^{-1/2}).$$

Thus, $\widehat{\psi}_n^{\mathrm{dml}}$ is an asymptotically linear estimator of $\Psi(P_0)$ with its influence function being the EIF of Ψ. It follows that $\widehat{\psi}_n^{\mathrm{dml}}$ is a regular and efficient estimator of Ψ. □

C.4.1 Proof of Theorem 4.15

Proof of Theorem 4.15. By Conditions A2 and A3, we have

$$\begin{aligned}\Psi_{\mathcal{H}'}(P_0) - \Psi_{\mathcal{H}}(P_0) &= \psi_0(\theta_{0,\mathcal{H}'}) - \psi_0(\theta_0) \\ &= \dot{\psi}_0(\theta_0)(\theta_{0,\mathcal{H}'} - \theta_0) + O\left(\|\theta_{0,\mathcal{H}'} - \theta_0\|_{\mathcal{H}}^2\right) \\ &= \partial_\theta^2 L_0(\theta_0, \eta_0)(\alpha_0, \theta_{0,\mathcal{H}'} - \theta_0) + O\left(\|\theta_{0,\mathcal{H}'} - \theta_0\|_{\mathcal{H}}^2\right),\end{aligned} \tag{C.15}$$

where the final equality uses the Riesz representation property of α_0, and the $O(\|\theta_{0,\mathcal{H}'} - \theta_0\|_{\mathcal{H}}^2)$ term is absent when the functional is linear.

Next, the first-order conditions defining the risk minimizer $\theta_{0,\mathcal{H}'}$ imply that

$$\partial_\theta L_0(\theta_{0,\mathcal{H}'}, \eta_0)(h) = 0 \quad \text{for all } h \in \mathcal{H}'.$$

Applying Condition A3 and using the first-order condition at θ_0, we find, for all $h \in \mathcal{H}'$ with $\rho_{\mathcal{H}}(h) < \infty$,

$$0 = \partial_\theta L_0(\theta_{0,\mathcal{H}'}, \eta_0)(h) = \partial_\theta^2 L_0(\theta_0, \eta_0)(\theta_{0,\mathcal{H}'} - \theta_0, h) + O\left(\rho_{\mathcal{H}}(h)\|\theta_{0,\mathcal{H}'} - \theta_0\|_{\mathcal{H}}^2\right).$$

The displayed remainder is absent when the risk is quadratic. Taking $h = \alpha_{0,\mathcal{H}'}$, whose auxiliary norm is finite by assumption, we find that

$$\partial_\theta^2 L_0(\theta_0, \eta_0)(\theta_{0,\mathcal{H}'} - \theta_0, \alpha_{0,\mathcal{H}'}) = O\left(\|\theta_{0,\mathcal{H}'} - \theta_0\|_{\mathcal{H}}^2\right).$$

By symmetry of $\partial_\theta^2 L_0(\theta_0, \eta_0)$, combining the above expression with the final equality in (C.15) gives

$$\Psi_{\mathcal{H}'}(P_0) - \Psi_{\mathcal{H}}(P_0) = \partial_\theta^2 L_0(\theta_0, \eta_0)(\alpha_0 - \alpha_{0,\mathcal{H}'}, \theta_{0,\mathcal{H}'} - \theta_0) + O\left(\|\theta_{0,\mathcal{H}'} - \theta_0\|_{\mathcal{H}}^2\right),$$

where the $O(\|\theta_{0,\mathcal{H}'} - \theta_0\|_{\mathcal{H}}^2)$ term is absent when the functional is linear and the risk is quadratic. This proves the claim. □

BIBLIOGRAPHY

Hervé Abdi, Dominique Valentin, and Betty Edelman. *Neural networks*. Number 124 in Sage University Papers Series on Quantitative Applications in the Social Sciences. Sage, Thousand Oaks, CA, 1999.

Alen Alexanderian. Optimization in infinite-dimensional hilbert spaces. *North Carolina State University, Raleigh, NC, USA*, 2019.

Martin Anthony and Peter L Bartlett. *Neural network learning: Theoretical foundations*. cambridge university press, 2009.

Anestis Antoniadis. Wavelet methods in statistics: some recent developments and their applications. *Statistics Surveys*, 1(none), jan 2007. doi: 10.1214/07-ss014. URL https://doi.org/10.1214%2F07-ss014.

Facundo Argañaraz. Automatic debiased machine learning of structural parameters with general conditional moments. *arXiv preprint arXiv:2512.08423*, 2025.

Susan Athey and Stefan Wager. Policy learning with observational data. *Econometrica*, 89(1):133–161, 2021.

Susan Athey, Guido W Imbens, and Stefan Wager. Approximate residual balancing: debiased inference of average treatment effects in high dimensions. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 80(4):597–623, 2018.

Vahe Avagyan and Stijn Vansteelandt. Honest data-adaptive inference for the average treatment effect under model misspecification using penalised bias-reduced double-robust estimation. *arXiv preprint arXiv:1708.03787*, 2017.

Andrii Babii. Identification and estimation in the functional linear instrumental regression. Technical report, Tech. Rep., UNC Working paper, 2017.

Andrii Babii and Jean-Pierre Florens. Is completeness necessary? estimation in non-identified linear models. *Econometric Theory*, pages 1–38, 2020.

Sivaraman Balakrishnan, Edward H Kennedy, and Larry Wasserman. The fundamental limits of structure-agnostic functional estimation. *arXiv preprint arXiv:2305.04116*, 2023.

Daniele Ballinari and Nora Bearth. Improving the finite sample performance of double/debiased machine learning with propensity score calibration. *arXiv preprint arXiv:2409.04874*, 2024.

Heejung Bang and James M Robins. Doubly robust estimation in missing data and causal inference models. *Biometrics*, 61(4):962–973, 2005.

Richard E Barlow and Hugh D Brunk. The isotonic regression problem and its dual. *Journal of the American Statistical Association*, 67(337):140–147, 1972.

Adam Baybutt and Manu Navjeevan. Doubly-robust inference for conditional average treatment effects with high-dimensional controls. *Journal of Econometrics*, 253:106180, 2026.

Antonio Bella, Cèsar Ferri, José Hernández-Orallo, and María José Ramírez-Quintana. Calibration of machine learning models. In *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*, pages 128–146. IGI Global, 2010.

Alexandre Belloni, Victor Chernozhukov, Denis Chetverikov, and Kengo Kato. Some new asymptotic theory for least squares series: Pointwise and uniform results. *Journal of Econometrics*, 186(2):345–366, 2015. doi: 10.1016/j.jeconom.2015.02.014.

Eli Ben-Michael, Avi Feller, David A Hirshberg, and José R Zubizarreta. The balancing act in causal inference. *arXiv preprint arXiv:2110.14831*, 2021.

David Benkeser and Nima S Hejazi. Doubly-robust inference in r using drtmle. *Observational Studies*, 9(2):43–78, 2023.

David Benkeser and Mark van der Laan. The highly adaptive lasso estimator. *International Conference on Data Science and Advanced Analytics*, pages 689–696, 2016. doi: 10.1109/DSAA.2016.93.

David Benkeser and Mark Van Der Laan. The highly adaptive lasso estimator. In *2016 IEEE international conference on data science and advanced analytics (DSAA)*, pages 689–696. IEEE, 2016.

David Benkeser, Marco Carone, MJ Van Der Laan, and Peter B Gilbert. Doubly robust nonparametric inference on the average treatment effect. *Biometrika*, 104(4): 863–880, 2017.

David Benkeser, Weixin Cai, and Mark Laan. A nonparametric super-efficient estimator of the average treatment effect. *Statistical Science*, 35:484–495, 08 2020. doi: 10.1214/19-STS735.

Andrew Bennett, Nathan Kallus, Xiaojie Mao, Whitney Newey, Vasilis Syrgkanis, and Masatoshi Uehara. Minimax instrumental variable regression and l_2 convergence guarantees without identification or closedness. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 2291–2318. PMLR, 2023a.

Andrew Bennett, Nathan Kallus, Xiaojie Mao, Whitney Newey, Vasilis Syrgkanis, and Masatoshi Uehara. Source condition double robust inference on functionals of inverse problems. *arXiv preprint arXiv:2307.13793*, 2023b.

Andrew Bennett, Nathan Kallus, Xiaojie Mao, Whitney K Newey, Vasilis Syrgkanis, and Masatoshi Uehara. Inference on strongly identified functionals of weakly identified functions. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, page qkaf075, 2025.

Dimitri Bertsekas, Angelia Nedic, and Asuman Ozdaglar. *Convex analysis and optimization*, volume 1. Athena Scientific, 2003.

Gérard Biau. Analysis of a random forests model. *The Journal of Machine Learning Research*, 13(1):1063–1095, 2012.

Gérard Biau, Luc Devroye, and Gäbor Lugosi. Consistency of random forests and other averaging classifiers. *Journal of Machine Learning Research*, 9(9), 2008.

Aurelien Bibaut, Maya Petersen, Nikos Vlassis, Maria Dimakopoulou, and Mark van der Laan. Sequential causal inference in a single world of connected units. *arXiv preprint arXiv:2101.07380*, 2021.

Peter J Bickel, Chris AJ Klaassen, YA'Acov Ritov, and JA Wellner. *Efficient and adaptive estimation for semiparametric models*, volume 4. Johns Hopkins University Press Baltimore, 1993.

Matteo Bonvini, Edward H Kennedy, Oliver Dukes, and Sivaraman Balakrishnan. Doubly-robust inference and optimality in structure-agnostic models with smoothness. *arXiv preprint arXiv:2405.08525*, 2024.

Jelena Bradic, Victor Chernozhukov, Whitney K Newey, and Yinchu Zhu. Minimax semiparametric learning with approximate sparsity. *arXiv preprint arXiv:1912.12213*, 2019a.

Jelena Bradic, Stefan Wager, and Yinchu Zhu. Sparsity double robust inference of average treatment effects. *arXiv preprint arXiv:1905.00744*, 2019b.

Jelena Bradic, Weijie Ji, and Yuqian Zhang. High-dimensional inference for dynamic treatment effects. *The Annals of Statistics*, 52(2):415–440, 2024.

L Breiman, JH Friedman, RA Olshen, and CJ Stone. *Classification and Regression Trees.* Belmont: Wadsworth, Belmont, CA, 1984.

Leo Breiman. Random forests. *Machine learning*, 45:5–32, 2001.

Jeanne Brooks-Gunn, Fong-ruey Liaw, and Pamela Kato Klebanov. Effects of early intervention on cognitive function of low birth weight preterm infants. *The Journal of pediatrics*, 120(3):350–359, 1992.

David Bruns-Smith, Oliver Dukes, Avi Feller, and Elizabeth L. Ogburn. Augmented balancing weights as linear regression. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, page qkaf019, 2025.

Marco Carone, Alexander R. Luedtke, and Mark J. van der Laan. Toward computerized efficient estimation in infinite-dimensional models. *Journal of the American Statistical Association*, 114(527):1174–1190, 2019. doi: 10.1080/01621459.2018.1482752. URL https://doi.org/10.1080/01621459.2018.1482752. PMID: 32405108.

Marine Carrasco, Jean-Pierre Florens, and Eric Renault. Linear inverse problems in structural econometrics estimation based on spectral decomposition and regularization. *Handbook of econometrics*, 6:5633–5751, 2007.

Gary Chamberlain. Asymptotic efficiency in estimation with conditional moment restrictions. *Journal of econometrics*, 34(3):305–334, 1987.

Sabyasachi Chatterjee, Adityanand Guntuboyina, and Bodhisattva Sen. Improved risk bounds in isotonic regression. *arXiv preprint arXiv:1311.3765*, 2013.

Shuai Chen, Lu Tian, Tianxi Cai, and Menggang Yu. A general statistical framework for subgroup identification and comparative treatment scoring. *Biometrics*, 73(4): 1199–1209, 2017.

Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, Rory Mitchell, Ignacio Cano, Tianyi Zhou, et al. Xgboost: extreme gradient boosting. *R package version 0.4-2*, 1(4):1–4, 2015.

Xiaohong Chen. Large sample sieve estimation of semi-nonparametric models. In James J. Heckman and Edward E. Leamer, editors, *Handbook of Econometrics*, volume 6B, chapter 76. Elsevier, 2007.

Xiaohong Chen and Timothy Christensen. Optimal uniform convergence rates for sieve nonparametric instrumental variables regression. *arXiv preprint arXiv:1311.0412*, 2013.

Xiaohong Chen and Zhipeng Liao. Sieve m inference on irregular parameters. *Journal of Econometrics*, 182(1):70–86, 2014. ISSN 0304-4076. doi: https://doi.org/10.1016/j.jeconom.2014.04.009. URL https://www.sciencedirect.com/science/article/pii/S0304407614000682. Causality, Prediction, and Specification Analysis: Recent Advances and Future Directions.

Xiaohong Chen, Zhipeng Liao, and Yixiao Sun. Sieve inference on possibly misspecified semi-nonparametric time series models. *Journal of Econometrics*, 178:639–658, 2014.

Yuxi Chen, Edward H. Kennedy, and Sivaraman Balakrishnan. On the equivalence between neyman orthogonality and pathwise differentiability, 2026. URL https://arxiv.org/abs/2603.15817.

Nicolas Chenavier, Norbert Henze, and Moritz Otto. Limit laws for large th-nearest neighbor balls. *Journal of Applied Probability*, 59(3):880–894, 2022.

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

Victor Chernozhukov, Whitney K Newey, and James Robins. Double/de-biased machine learning using regularized riesz representers. Technical report, cemmap working paper, 2018b.

Victor Chernozhukov, Carlos Cinelli, Whitney K Newey, Amit Sharma, and Vasilis Syrgkanis. Long story short: Omitted variable bias in causal machine learning. *The Review of Economics and Statistics*, pages 1–45, 2021a.

Victor Chernozhukov, Whitney K Newey, Victor Quintas-Martinez, and Vasilis Syrgkanis. Automatic debiased machine learning via neural nets for generalized linear regression. *arXiv preprint arXiv:2104.14737*, 2021b.

Victor Chernozhukov, Whitney Newey, Victor M Quintas-Martinez, and Vasilis Syrgkanis. Riesznet and forestriesz: Automatic debiased machine learning with neural nets and random forests. In *International Conference on Machine Learning*, pages 3901–3914. PMLR, 2022a.

Victor Chernozhukov, Whitney Newey, Rahul Singh, and Vasilis Syrgkanis. Automatic debiased machine learning for dynamic treatment effects and general nested functionals, 2022b.

Victor Chernozhukov, Whitney K Newey, and Rahul Singh. Automatic debiased machine learning of causal and structural effects. *Econometrica*, 90(3):967–1027, 2022c.

Victor Chernozhukov, Whitney K Newey, and Rahul Singh. Debiased machine learning of global and local parameters using regularized riesz representers. *The Econometrics Journal*, 25(3):576–601, 2022d.

Victor Chernozhukov, Michael Newey, Whitney K Newey, Rahul Singh, and Vasilis Srygkanis. Automatic debiased machine learning for covariate shifts. *arXiv preprint arXiv:2307.04527*, 2023.

Victor Chernozhukov, Whitney K Newey, Victor Quintas-Martinez, and Vasilis Syrgkanis. Automatic debiased machine learning via riesz regression. *arXiv e-prints*, pages arXiv–2104, 2024a.

Victor Chernozhukov, Whitney K Newey, and Vasilis Syrgkanis. Conditional influence functions. *arXiv preprint arXiv:2412.18080*, 2024b.

Victor Chernozhukov, Whitney K Newey, Rahul Singh, and Vasilis Syrgkanis. Adversarial estimation of riesz representers. *Journal of the American Statistical Association*, pages 1–12, 2026.

Brian Cho, Yaroslav Mukhin, Kyra Gan, and Ivana Malenica. Kernel debiased plug-in estimation: Simultaneous, automated debiasing without influence functions for many target parameters. *Proceedings of machine learning research*, 235:8534, 2024.

Fernando Cobos, Thomas Kühn, and Winfried Sickel. Optimal approximation of multivariate periodic sobolev functions in the sup-norm. *Journal of Functional Analysis*, 270(11):4196–4212, 2016.

Yifan Cui, Hongming Pu, Xu Shi, Wang Miao, and Eric Tchetgen Tchetgen. Semiparametric proximal causal inference. *Journal of the American Statistical Association*, 119(546):1348–1359, 2024.

Alicia Curth and Mihaela van der Schaar. Nonparametric estimation of heterogeneous treatment effects: From theory to learning algorithms. In *International Conference on Artificial Intelligence and Statistics*, pages 1810–1818. PMLR, 2021.

Alicia Curth, Ahmed M Alaa, and Mihaela van der Schaar. Estimating structural target functions using machine learning and influence functions. *arXiv preprint arXiv:2008.06461*, 2020.

Alexander D'Amour and Alexander Franks. Deconfounding scores: Feature representations for causal effect estimation with weak overlap. *arXiv preprint arXiv:2104.05762*, 2021.

Zhun Deng, Cynthia Dwork, and Linjun Zhang. Happymap: A generalized multicalibration method. *arXiv preprint arXiv:2303.04379*, 2023.

Shachi Deshpande and Volodymyr Kuleshov. Calibrated and conformal propensity scores for causal effect estimation. In *The 40th Conference on Uncertainty in Artificial Intelligence*, 2023.

Jean-Claude Deville and Carl-Erik Särndal. Calibration estimators in survey sampling. *Journal of the American statistical Association*, 87(418):376–382, 1992.

Iván Díaz. Statistical inference for data-adaptive doubly robust estimators with survival outcomes. *Statistics in medicine*, 38(15):2735–2748, 2019.

Iván Díaz. Non-agency interventions for causal mediation in the presence of intermediate confounding. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86(2):435–460, 2024.

Iván Díaz and Mark J van der Laan. Targeted data adaptive estimation of the causal dose–response curve. *Journal of Causal Inference*, 1(2):171–192, 2013.

Iván Díaz and Mark J van der Laan. Doubly robust inference for targeted minimum loss–based estimation in randomized trials with missing outcome data. *Statistics in medicine*, 36(24):3807–3819, 2017.

Oliver Dukes, Stijn Vansteelandt, and David Whitney. On doubly robust inference for double machine learning in semiparametric regression. *Journal of Machine Learning Research*, 25(279):1–46, 2024.

Bradley Efron and Robert J Tibshirani. *An introduction to the bootstrap*. CRC press, 1994.

Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*. SIAM, 1999.

Billy Fang, Adityanand Guntuboyina, and Bodhisattva Sen. Multivariate extensions of isotonic regression and total variation denoising via entire monotonicity and hardy–krause variation. *The Annals of Statistics*, 49(2):769–792, 2021.

M. Farrell, Tengyuan Liang, and S. Misra. Deep neural networks for estimation and inference: Application to causal effects and other semiparametric estimands. *ArXiv*, abs/1809.09953, 2018.

Hamza Fawzi. Topics in convex optimisation (l16). *Mathematical Tripos Part III Guide to Courses 2016-2017*, page 71, 2017.

Aaron Fisher. Inverse-variance weighting for estimation of heterogeneous treatment effects. In *Forty-first International Conference on Machine Learning*, 2024.

Evelyn Fix and Joseph Lawson Hodges. Discriminatory analysis. nonparametric discrimination: Consistency properties. *International Statistical Review/Revue Internationale de Statistique*, 57(3):238–247, 1989.

Dylan J Foster and Vasilis Syrgkanis. Orthogonal statistical learning. *The Annals of Statistics*, 51(3):879–908, 2023.

Constantine E Frangakis, Tianchen Qian, Zhenke Wu, and Ivan Diaz. Deductive derivation and turing-computerization of semiparametric efficient estimation. *Biometrics*, 71(4):867–874, 2015.

Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55 (1):119–139, 1997. ISSN 0022-0000. doi: https://doi.org/10.1006/jcss.1997.1504. URL https://www.sciencedirect.com/science/article/pii/S002200009791504X.

Jerome H Friedman. Multivariate adaptive regression splines. *The annals of statistics*, 19(1):1–67, 1991.

AmirEmad Ghassami, James M Robins, and Andrea Rotnitzky. Debiased ill-posed regression. *arXiv preprint arXiv:2505.20787*, 2025.

Satyajit Ghosh and Zhiqiang Tan. Doubly robust semiparametric inference using regularized calibrated estimation with high-dimensional data. *Bernoulli*, 28(3):1675–1703, 2022.

William J. Gordon and Richard F. Riesenfeld. B-spline curves and surfaces. *Computer Aided Geometric Design*, pages 95–126, 1974.

Ellen Graham, Marco Carone, and Andrea Rotnitzky. Towards a unified theory for semiparametric data fusion with individual-level data. *arXiv preprint arXiv:2409.09973*, 2024.

Piet Groeneboom and HP Lopuhaa. Isotonic estimators of monotone densities and distribution functions: basic facts. *Statistica Neerlandica*, 47(3):175–183, 1993.

Susan Gruber, Rachael V Phillips, Hana Lee, and Mark J van der Laan. Data-adaptive selection of the propensity score truncation level for inverse-probability–weighted and targeted maximum likelihood estimators of marginal point treatment effects. *American Journal of Epidemiology*, 191(9):1640–1651, 2022.

Adityanand Guntuboyina and Bodhisattva Sen. Nonparametric shape-restricted regression. *Statistical Science*, 33(4):568–594, 2018.

Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q Weinberger. On calibration of modern neural networks. In *International conference on machine learning*, pages 1321–1330. PMLR, 2017.

Chirag Gupta and Aaditya Ramdas. Distribution-free calibration guarantees for histogram binning without sample splitting. In *International Conference on Machine Learning*, pages 3942–3952. PMLR, 2021.

Chirag Gupta, Aleksandr Podkopaev, and Aaditya Ramdas. Distribution-free binary classification: prediction sets, confidence intervals and calibration. *Advances in Neural Information Processing Systems*, 33:3711–3723, 2020.

Rom Gutman, Ehud Karavani, and Yishai Shimoni. Propensity score models are better when post-calibrated. *arXiv preprint arXiv:2211.01221*, 2022.

László Györfi, Michael Kohler, Adam Krzyżak, and Harro Walk. *A distribution-free theory of nonparametric regression*. Springer, 2002.

Jinyong Hahn, Zhipeng Liao, and Geert Ridder. Nonparametric two-step sieve m estimation and inference. *Econometric Theory*, 34(6):1281–1324, 2018. doi: 10.1017/S0266466618000014.

P Richard Hahn, Vincent Dorie, and Jared S Murray. Atlantic causal inference conference (acic) data analysis challenge 2017. *arXiv preprint arXiv:1905.09515*, 2019.

Jaroslav Hájek. Local asymptotic minimax and admissibility in estimation. In *Proceedings of the sixth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 175–194, 1972.

Rafail Z Hasminskii and Ildar A Ibragimov. On the nonparametric estimation of functionals. In *Proceedings of the Second Prague Symposium on Asymptotic Statistics*, volume 473, pages 474–482. North-Holland Amsterdam, 1979.

Trevor Hastie and Maintainer Trevor Hastie. Package 'gam'. *R package version*, pages 90124–3, 2015.

Trevor Hastie and Robert Tibshirani. Generalized additive models: some applications. *Journal of the American Statistical Association*, 82(398):371–386, 1987.

Trevor J Hastie and Robert J Tibshirani. *Generalized additive models*, volume 43. CRC press, 1990.

Nima S Hejazi, Jeremy R Coyle, and Mark J van der Laan. hal9001: Scalable highly adaptive lasso regression in R. *Journal of Open Source Software*, 2020. doi: 10.21105/joss.02526. URL https://doi.org/10.21105/joss.02526.

Jennifer L Hill. Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, 20(1):217–240, 2011.

Oliver J Hines and Caleb H Miles. Learning density ratios in causal inference using bregman-riesz regression. *arXiv preprint arXiv:2510.16127*, 2025.

David A. Hirshberg and Stefan Wager. Augmented minimax linear estimation. *The Annals of Statistics*, 49(6):3206 – 3227, 2021. doi: 10.1214/21-AOS2080. URL https://doi.org/10.1214/21-AOS2080.

David P Hofmeyr. Fast kernel smoothing in r with applications to projection pursuit. *arXiv preprint arXiv:2001.02225*, 2020.

Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5):359–366, 1989.

Jianhua Z Huang. Asymptotics for polynomial spline regression under weak conditions. *Statistics & probability letters*, 65(3):207–216, 2003.

David Hubbard, Benoit Rostykus, Yves Raimond, and Tony Jebara. Beta survival models. In *Survival prediction-algorithms, challenges and applications*, pages 22–39. PMLR, 2021.

Hidehiko Ichimura and Whitney K Newey. The influence function of semiparametric estimators. *Quantitative Economics*, 13(1):29–61, 2022.

Kosuke Imai and Marc Ratkovic. Covariate balancing propensity score. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, pages 243–263, 2014.

Dunham Jackson. On approximation by trigonometric sums and polynomials. *Transactions of the American Mathematical Society*, 13:491–515, 1912.

Binyan Jiang, Rui Song, Jialiang Li, and Donglin Zeng. Entropy learning for dynamic treatment regimes. *Statistica Sinica*, 29:1633–1710, 01 2019. doi: 10.5705/ss.202018.0076.

Jikai Jin and Vasilis Syrgkanis. Structure-agnostic optimality of doubly robust learning for treatment effect estimation. *arXiv preprint arXiv:2402.14264*, 2024.

Jikai Jin and Vasilis Syrgkanis. Sharp structure-agnostic lower bounds for general functional estimation. *arXiv preprint arXiv:2512.17341*, 2025.

Alexander I Jordan, Anja Mühlemann, and Johanna F Ziegel. Characterizing the optimal solutions to the isotonic regression problem for identifiable functionals. *Annals of the Institute of Statistical Mathematics*, 74(3):489–514, 2022a.

Michael Jordan, Yixin Wang, and Angela Zhou. Empirical gateaux derivatives for causal inference. *Advances in Neural Information Processing Systems*, 35:8512–8525, 2022b.

Nathan Kallus and Masatoshi Uehara. Double reinforcement learning for efficient off-policy evaluation in markov decision processes. *Journal of Machine Learning Research*, 21(167):1–63, 2020.

Nathan Kallus and Masatoshi Uehara. Efficiently breaking the curse of horizon in off-policy evaluation with double reinforcement learning. *Operations Research*, 70(6): 3282–3302, 2022.

J. D. Y. Kang and J. L. Schafer. Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical Science*, 22:523—-39, 2007.

Göran Kauermann and Raymond J Carroll. The sandwich variance estimator: Efficiency properties and coverage probability of confidence intervals. Technical Report Discussion Paper 189, Collaborative Research Center 386, Ludwig-Maximilians-Universität München, 2000.

Edward H. Kennedy. Towards optimal doubly robust estimation of heterogeneous causal effects. *Electronic Journal of Statistics*, 17(2):3008–3049, 2023a. doi: 10.1214/23-EJS2157.

Edward H Kennedy. Towards optimal doubly robust estimation of heterogeneous causal effects. *Electronic Journal of Statistics*, 17(2):3008–3049, 2023b.

Edward H Kennedy, Zongming Ma, Matthew D McHugh, and Dylan S Small. 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.

Edward H. Kennedy, Sivaraman Balakrishnan, James M. Robins, and Larry Wasserman. Minimax rates for heterogeneous causal effect estimation. *The Annals of Statistics*, 52(2):793–816, 2024. doi: 10.1214/24-AOS2369.

Michael P Kim, Amirata Ghorbani, and James Zou. Multiaccuracy: Black-box post-processing for fairness in classification. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pages 247–254, 2019.

Sven Klaassen, Jan Rabenseifner, Jannis Kueck, and Philipp Bach. Calibration strategies for robust causal estimation: Theoretical and empirical insights on propensity score-based estimators. *arXiv preprint arXiv:2503.17290*, 2025.

Mark A Kon and Louise Arakelian Raphael. Sup-norm convergence rates of wavelet expansions in besov spaces. *Applicable Mathematics*, pages 193–203, 2002.

Sören R Künzel, Jasjeet S Sekhon, Peter J Bickel, and Bin Yu. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the national academy of sciences*, 116(10):4156–4165, 2019.

Mark J. van der Laan and Daniel Rubin. Targeted maximum likelihood learning. *The International Journal of Biostatistics*, 2(1), 2006. doi: 10.2202/1557-4679.1043.

Mark J van der Laan, Sandrine Dudoit, and Aad W van der Vaart. The cross-validated adaptive epsilon-net estimator. *Statistics & Decisions*, 24(3):373–395, 2006.

Robert J LaLonde. Evaluating the econometric evaluations of training programs with experimental data. *The American economic review*, pages 604–620, 1986.

Lucien Le Cam. *Asymptotic methods in statistical decision theory*. Springer Science & Business Media, 2012.

Kaitlyn J Lee and Alejandro Schuler. Rieszboost: Gradient boosting for riesz regression. *arXiv preprint arXiv:2501.04871*, 2025.

Hannes Leeb and Benedikt M. Pötscher. Model selection and inference: Facts and fiction. *Econometric Theory*, 21(1):21–59, 2005. ISSN 02664666, 14694360. URL http://www.jstor.org/stable/3533623.

Liu Leqi and Edward H Kennedy. Median optimal treatment regimes. *arXiv preprint arXiv:2103.01802*, 2021.

Boris Yakovlevich Levit. On efficiency of a class of non-parametric estimates. *Teoriya Veroyatnostei i ee Primeneniya*, 20(4):738–754, 1975.

Fan Li, Kari Lock Morgan, and Alan M Zaslavsky. Balancing covariates via propensity score weighting. *Journal of the American Statistical Association*, 113(521):390–400, 2018.

Fan Li, Laine E Thomas, and Fan Li. Addressing extreme propensity scores via the overlap weights. *American journal of epidemiology*, 188(1):250–257, 2019.

Lingling Li, Eric Tchetgen Tchetgen, Aad van der Vaart, and James M Robins. Higher order inference on a treatment effect under low regularity conditions. *Statistics & probability letters*, 81(7):821–828, 2011.

Yi Li, Sky Qiu, Zeyi Wang, and Mark van der Laan. Regularized targeted maximum likelihood estimation in highly adaptive lasso implied working models. *arXiv preprint arXiv:2506.17214*, 2025.

Sarah Lichtenstein, Baruch Fischhoff, and Lawrence D Phillips. Calibration of probabilities: The state of the art. In *Decision Making and Change in Human Affairs: Proceedings of the Fifth Research Conference on Subjective Probability, Utility, and Decision Making, Darmstadt, 1–4 September, 1975*, pages 275–324. Springer, 1977.

Lin Liu, Rajarshi Mukherjee, Whitney K Newey, and James M Robins. Semiparametric efficient empirical higher order influence function estimators. *arXiv preprint arXiv:1705.07577*, 2017.

Lin Liu, Xinbo Wang, and Yuhao Wang. Root-n consistent semiparametric learning with high-dimensional nuisance functions under minimal sparsity. *arXiv preprint arXiv:2305.04174*, 2023.

Christos Louizos, Uri Shalit, Joris M Mooij, David Sontag, Richard Zemel, and Max Welling. Causal effect inference with deep latent-variable models. *Advances in neural information processing systems*, 30, 2017.

Alex Luedtke. Simplifying debiased inference via automatic differentiation and probabilistic programming. *arXiv preprint arXiv:2405.08675*, 2024.

Alex Luedtke and Incheoul Chung. One-step estimation of differentiable hilbert-valued parameters. *The Annals of Statistics*, 52(4):1534–1563, 2024. doi: 10.1214/24-AOS2403.

Alexander Luedtke and Mark Laan. Super-learning of an optimal dynamic treatment rule. *The international journal of biostatistics*, 12:305–332, 05 2016. doi: 10.1515/ijb-2015-0052.

Alexander R Luedtke, Marco Carone, and Mark J van der Laan. Discussion of "deductive derivation and turing-computerization of semiparametric efficient estimation" by frangakis et al. *Biometrics*, 71(4):875–879, 2015.

Alexander R Luedtke, Oleg Sofrygin, Mark J van der Laan, and Marco Carone. Sequential double robustness in right-censored longitudinal models. *arXiv preprint arXiv:1705.02459*, 2017.

Enno Mammen and Sara van de Geer. Locally adaptive regression splines. *The Annals of Statistics*, 25(1):387 – 413, 1997. doi: 10.1214/aos/1034276635. URL https://doi.org/10.1214/aos/1034276635.

Alec McClean, Zach Branson, and Edward H Kennedy. Nonparametric estimation of conditional incremental effects. *arXiv preprint arXiv:2212.03578*, 2022.

Alec McClean, Sivaraman Balakrishnan, Edward H Kennedy, and Larry Wasserman. Double cross-fit doubly robust estimators: Beyond series regression. *arXiv preprint arXiv:2403.15175*, 2024.

Sean McGrath and Rajarshi Mukherjee. Nuisance function tuning and sample splitting for optimal doubly robust estimation. *arXiv preprint arXiv:2212.14857*, 2022.

Shahar Mendelson and Joseph Neeman. Regularization in kernel learning. *The Annals of Statistics*, 38(1):526–565, 2010. doi: 10.1214/09-AOS728.

Maintainer Stephen Milborrow. Package 'earth'. *R Software package*, 2019.

Pawel Morzywolek, Johan Decruyenaere, and Stijn Vansteelandt. On a general class of orthogonal learners for the estimation of heterogeneous treatment effects. *arXiv preprint arXiv:2303.12687*, 2023.

Jaouad Mourtada, Stéphane Gaïffas, and Erwan Scornet. Minimax optimal rates for mondrian trees and forests. *The Annals of Statistics*, 48(4):2253–2276, 2020. doi: 10.1214/19-AOS1886.

Susan A Murphy and Aad W Van der Vaart. On profile likelihood. *Journal of the American Statistical Association*, 95(450):449–465, 2000.

Elizbar A Nadaraya. On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142, 1964.

Brady Neal, Chin-Wei Huang, and Sunand Raghupathi. Realcause: Realistic causal inference benchmarking. *arXiv preprint arXiv:2011.15007*, 2020.

Denis Nekipelov, Vira Semenova, and Vasilis Syrgkanis. Regularised orthogonal machine learning for nonlinear semiparametric models. *The Econometrics Journal*, 25(1): 233–255, 2022.

Whitney K. Newey. The asymptotic variance of semiparametric estimators. *Econometrica*, 62(6):1349–1382, 1994. doi: 10.2307/2951752.

Whitney K Newey and James R Robins. Cross-fitting and fast remainder rates for semiparametric estimation. *arXiv preprint arXiv:1801.09138*, 2018.

Whitney K Newey, Fushing Hsieh, and James Robins. Undersmoothing and bias corrected functional estimation. Technical Report Working Paper 98-17, Massachusetts Institute of Technology, Department of Economics, Cambridge, MA, 1998.

Richard Nickl and Benedikt M Pötscher. Bracketing metric entropy rates and empirical central limit theorems for function classes of besov-and sobolev-type. *Journal of Theoretical Probability*, 20(2):177–199, 2007.

Alexandru Niculescu-Mizil and Rich Caruana. Obtaining calibrated probabilities from boosting. In *UAI*, volume 5, pages 413–20, 2005a.

Alexandru Niculescu-Mizil and Rich Caruana. Predicting good probabilities with supervised learning. In *Proceedings of the 22nd international conference on Machine learning*, pages 625–632, 2005b.

Xinkun Nie and Stefan Wager. Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, 108(2):299–319, 2021.

Antonio Olivas-Martinez and Andrea Rotnitzky. Source-condition analysis of kernel adversarial estimators. *arXiv preprint arXiv:2508.17181*, 2025.

J Pfanzagl and W Wefelmeyer. Contributions to a general asymptotic statistical theory. *Statistics & Risk Modeling*, 3(3-4):379–388, 1985.

Noel Pimentel, Alejandro Schuler, and Mark van der Laan. Score-preserving targeted maximum likelihood estimation. *arXiv preprint arXiv:2502.00200*, 2025.

Kristin E Porter, Susan Gruber, Mark J Van Der Laan, and Jasjeet S Sekhon. The relative performance of targeted maximum likelihood estimators. *The international journal of biostatistics*, 7(1):000010220215574679 1308, 2011.

Hongxiang Qiu, Alex Luedtke, and Mark van der Laan. Contribution to discussion of "entropy learning for dynamic treatment regimes" by jiang b, song r, li j, zeng d. statistica sinica. *Statistica Sinica*, 29(4):1666–1678, 2019.

Hongxiang Qiu, Alex Luedtke, and Marco Carone. Universal sieve-based strategies for efficient estimation using machine learning tools. *Bernoulli: official journal of the Bernoulli Society for Mathematical Statistics and Probability*, 27(4):2300, 2021.

Jan Rabenseifner, Sven Klaassen, Jannis Kueck, and Philipp Bach. Calibration strategies for robust causal estimation: Theoretical and empirical insights on propensity score based estimators. *arXiv preprint arXiv:2503.17290*, 2025.

Louis B Rall. *Automatic differentiation: Techniques and applications*. Springer, 1981.

Thomas S Richardson, James M Robins, and Linbo Wang. On modeling and estimation for the relative risk and risk difference. *Journal of the American Statistical Association*, 112(519):1121–1130, 2017.

James Robins and Andrea Rotnitzky. Estimation of treatment effects in randomised trials with non-compliance and a dichotomous outcome using structural mean models. *Biometrika*, 91(4):763–783, 2004.

James Robins, Lingling Li, Eric Tchetgen, Aad van der Vaart, et al. Higher order influence functions and minimax estimation of nonlinear functionals. In *Probability and statistics: essays in honor of David A. Freedman*, volume 2, pages 335–422. Institute of Mathematical Statistics, 2008.

James M. Robins. Optimal structural nested models for optimal sequential decisions. In D. Y. Lin and P. J. Heagerty, editors, *Proceedings of the Second Seattle Symposium in Biostatistics: Analysis of Correlated Data*, pages 189–326. Springer New York, New York, NY, 2004. ISBN 978-1-4419-9076-1. doi: 10.1007/978-1-4419-9076-1_11. URL https://doi.org/10.1007/978-1-4419-9076-1_11.

James M. Robins, Andrea Rotnitzky, and Lue Ping Zhao. Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association*, 89(427):846–866, 1994. doi: 10.1080/01621459.1994.10476818.

James M Robins, Andrea Rotnitzky, and Lue Ping Zhao. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the american statistical association*, 90(429):106–121, 1995.

James M Robins, Miguel Angel Hernan, and Babette Brumback. Marginal structural models and causal inference in epidemiology, 2000.

Peter M Robinson. Root-n-consistent semiparametric regression. *Econometrica: Journal of the Econometric Society*, pages 931–954, 1988.

Sherri Rose and Mark J van der Laan. A targeted maximum likelihood estimator for two-stage designs. *The international journal of biostatistics*, 7(1):17, 2011.

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

Aaron Roth. Uncertain: Modern topics in uncertainty estimation. *Unpublished Lecture Notes*, 11(30-31):4, 2022.

Andrea Rotnitzky, James Robins, and Lucia Babino. On the multiply robust estimation of the mean of the g-functional. *arXiv preprint arXiv:1705.08582*, 2017.

Andrea Rotnitzky, Ezequiel Smucler, and James M Robins. Characterization of parameters with a mixed bias property. *Biometrika*, 108(1):231–238, 2021.

Andrea Rotnitzky, Ezequiel Smucler, and James M Robins. A note on the relation between one–step, outcome regression and ipw–type estimators of parameters with the mixed bias property. *Statistics & Probability Letters*, page 110796, 2026.

HL Royden. Real analysis, the macmillan company. *New York*, 1963.

Daniel Rubin and Mark J van der Laan. Doubly robust censoring unbiased transformations. 2006.

Daniel Rubin and Mark J van der Laan. A doubly robust censoring unbiased transformation. *The international journal of biostatistics*, 3(1), 2007.

Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.

Donald B Rubin. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469):322–331, 2005.

Walter Rudin. Functional analysis 2nd ed. *International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York*, 1991.

Anton Schick. On asymptotically efficient estimation in semiparametric models. *The Annals of Statistics*, pages 1139–1151, 1986.

Alejandro Schuler, Yi Li, and Mark van der Laan. Lassoed tree boosting, 2023. URL https://arxiv.org/abs/2205.10697.

Rajen D Shah and Jonas Peters. The hardness of conditional independence testing and the generalised covariance measure. *The Annals of Statistics*, 48(3):1514–1538, 2020. doi: 10.1214/19-AOS1857.

Uri Shalit, Fredrik D Johansson, and David Sontag. Estimating individual treatment effect: generalization bounds and algorithms. In *International conference on machine learning*, pages 3076–3085. PMLR, 2017.

Anne Sheehy and Jon A Wellner. Uniform donsker classes of functions. *The Annals of Probability*, 20(4):1983–2030, 1992.

Xiaotong Shen. On methods of sieves and penalization. *The Annals of Statistics*, 25 (6):2555–2591, 1997.

Yishai Shimoni, Chen Yanover, Ehud Karavani, and Yaara Goldschmnidt. Benchmarking framework for performance-evaluation of causal inference analysis. *arXiv preprint arXiv:1802.05046*, 2018.

Toru Shirakawa, Yi Li, Yulun Wu, Sky Qiu, Yuxuan Li, Mingduo Zhao, Hiroyasu Iso, and Mark Van Der Laan. Longitudinal targeted minimum loss-based estimation with temporal-difference heterogeneous transformer. *Proceedings of machine learning research*, 235:45097, 2024.

Bonnie E Shook-Sa, Paul N Zivich, Chanhwa Lee, Keyi Xue, Rachael K Ross, Jessie K Edwards, Jeffrey SA Stringer, and Stephen R Cole. Double robust variance estimation with parametric working models. *Biometrics*, 81(2):ujaf054, 2025.

Ezequiel Smucler, Andrea Rotnitzky, and James M Robins. A unifying approach for doubly-robust l1-regularized estimation of causal contrasts. *arXiv preprint arXiv:1904.03737*, 2019.

Ezequiel Smucler, James M Robins, and Andrea Rotnitzky. On the asymptotic validity of confidence sets for linear functionals of solutions to integral equations. *Biometrika*, page asaf067, 2025.

Charles Stein. Efficient nonparametric testing and estimation. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 187–195, 1956.

Charles J Stone. Optimal global rates of convergence for nonparametric regression. *The annals of statistics*, pages 1040–1053, 1982.

Erik Sverdrup and Yifan Cui. Proximal causal learning of heterogeneous treatment effects. *arXiv preprint arXiv:2301.10913*, 2023.

Erik Sverdrup and Trevor Hastie. balnet: Pathwise estimation of covariate balancing propensity scores. *arXiv preprint arXiv:2602.18577*, 2026.

Vasilis Syrgkanis, Victor Lei, Miruna Oprescu, Maggie Hei, Keith Battocchi, and Greg Lewis. Machine learning estimation of heterogeneous treatment effects with instruments. *Advances in Neural Information Processing Systems*, 32, 2019.

Zhiqiang Tan. Model-assisted inference for treatment effects using regularized calibrated estimation with high-dimensional data. *The Annals of Statistics*, 48(2):811–837, 2020. doi: 10.1214/19-AOS1824.

Zhou Tang and Ted Westling. Consistency of the bootstrap for asymptotically linear estimators based on machine learning. *arXiv preprint arXiv:2404.03064*, 2024.

E.J. Tchetgen Tchetgen, J.M. Robins, and A. Rotnitzky. On doubly robust estimation in a semiparametric odds ratio model. *Biometrika*, 97:171–180, 2010.

Julie Tibshirani, Susan Athey, Rina Friedberg, Vitor Hadad, David Hirshberg, Luke Miner, Erik Sverdrup, Stefan Wager, Marvin Wright, and Maintainer Julie Tibshirani. Package 'grf', 2018.

Sara van de Geer. *Empirical Processes in M-estimation*, volume 6. Cambridge university press, 2000.

Sara van de Geer. On the uniform convergence of empirical norms and inner products, with application to causal inference. *Electronic Journal of Statistics*, 8:543–574, 2014.

Lars van der Laan and Ahmed Alaa. Generalized venn and venn-abers calibration with applications in conformal prediction. *arXiv preprint arXiv:2502.05676*, 2025.

Lars van der Laan and Ahmed M Alaa. Self-calibrating conformal prediction. *Advances in Neural Information Processing Systems*, 37:107138–107170, 2024.

Lars van der Laan, Marco Carone, Alex Luedtke, and Mark van der Laan. Adaptive debiased machine learning using data-driven model selection techniques. *arXiv preprint arXiv:2307.12544*, 2023.

Lars Van Der Laan, Ernesto Ulloa-Pérez, Marco Carone, and Alex Luedtke. Causal isotonic calibration for heterogeneous treatment effects. In *International Conference on Machine Learning*, pages 34831–34854. PMLR, 2023.

Lars van der Laan, Marco Carone, and Alex Luedtke. Combining t-learning and dr-learning: a framework for oracle-efficient estimation of causal contrasts. *arXiv preprint arXiv:2402.01972*, 2024a.

Lars van der Laan, Ziming Lin, Marco Carone, and Alex Luedtke. Stabilized inverse probability weighting via isotonic calibration. *arXiv preprint arXiv:2411.06342*, 2024b.

Lars van der Laan, Alex Luedtke, and Marco Carone. Doubly robust inference via calibration. *arXiv preprint arXiv:2411.02771*, 2024c.

Lars van der Laan, Aurélien Bibaut, and Nathan Kallus. Efficient inference for inverse reinforcement learning and dynamic discrete choice models. *CoRR*, abs/2512.24407, 2025a. doi: 10.48550/ARXIV.2512.24407. URL https://doi.org/10.48550/arXiv.2512.24407.

Lars van der Laan, Aurelien Bibaut, Nathan Kallus, and Alex Luedtke. Automatic debiased machine learning for smooth functionals of nonparametric M-estimands, 2025b. URL https://arxiv.org/abs/2501.11868.

Lars van der Laan, David Hubbard, Allen Tran, Nathan Kallus, and Aurélien Bibaut. Automatic double reinforcement learning in semiparametric markov decision processes with applications to long-term causal inference. *arXiv preprint arXiv:2501.06926*, 2025c.

Lars van der Laan, Nathan Kallus, and Aurélien Bibaut. Nonparametric instrumental variable inference with many weak instruments. *arXiv preprint arXiv:2505.07729*, 2025d.

M. van der Laan. A generally efficient targeted minimum loss based estimator. *Biostatistics Working Paper Series Working Paper 343.*, 2015.

Mark van der Laan. A generally efficient targeted minimum loss based estimator based on the highly adaptive lasso. *The international journal of biostatistics*, 13(2), 2017.

Mark van der Laan and Susan Gruber. One-step targeted minimum loss-based estimation based on universal least favorable one-dimensional submodels. *The international journal of biostatistics*, 12(1):351–378, 2016.

Mark van der Laan, David Benkeser, and Weixin Cai. Efficient estimation of pathwise differentiable target parameters with the undersmoothed highly adaptive lasso. *The International Journal of Biostatistics*, 2022. doi: doi:10.1515/ijb-2019-0092. URL https://doi.org/10.1515/ijb-2019-0092.

Mark van der Laan, Sky Qiu, Jens Magelund Tarp, and Lars van der Laan. Adaptivetmle for the average treatment effect based on randomized controlled trial augmented with real-world data. *Journal of Causal Inference*, 14(1):20240025, 2026.

Mark J. van der Laan. Targeted learning of an optimal dynamic treatment, and statistical inference for its mean outcome. *U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 317.*, 2013.

Mark J van der Laan. Targeted estimation of nuisance parameters to obtain valid statistical inference. *The international journal of biostatistics*, 10(1):29–57, 2014.

Mark J van der Laan and Sandrine Dudoit. Unified cross-validation methodology for selection among estimators and a general cross-validated adaptive epsilon-net estimator: Finite sample oracle inequalities and examples. Technical Report Working Paper 130, Division of Biostatistics, University of California, Berkeley, 2003. URL https://biostats.bepress.com/ucbbiostat/paper130/.

Mark J van der Laan and Susan Gruber. Targeted minimum loss based estimation of causal effects of multiple time point interventions. *The international journal of biostatistics*, 8(1), 2012.

Mark J. van der Laan and Alexander R. Luedtke. Targeted learning of the mean outcome under an optimal dynamic treatment rule. *Journal of Causal Inference*, 3(1): 61–95, 2015. doi: 10.1515/jci-2013-0022. URL https://doi.org/10.1515/jci-2013-0022.

Mark J. van der Laan and James M. Robins. *Unified Methods for Censored Longitudinal Data and Causality*. Springer Series in Statistics. Springer, New York, 2003. doi: 10.1007/978-0-387-21700-0.

Mark J van der Laan and Sherri Rose. *Targeted Learning: Causal Inference for Observational and Experimental Data*. Springer, New York, 2011.

Mark J. van der Laan and Sherri Rose. Why machine learning cannot ignore maximum likelihood estimation, 2021. URL https://arxiv.org/abs/2110.12112.

Mark J van der Laan, Alan Hubbard, and Nicholas P Jewell. Estimation of treatment effects in randomized trials with non-compliance and a dichotomous outcome. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 69(3):463–482, 2007.

Mark J van der Laan, Sherri Rose, Wenjing Zheng, and Mark J van der Laan. Cross-validated targeted minimum-loss-based estimation. *Targeted learning: causal inference for observational and experimental data*, pages 459–474, 2011.

Aad van der Vaart. Efficiency and hadamard differentiability. *Scandinavian Journal of Statistics*, 18(1):63–75, 1991. ISSN 03036898, 14679469. URL http://www.jstor.org/stable/4616189.

Aad van der Vaart and Jon A Wellner. A local maximal inequality under uniform entropy. *Electronic Journal of Statistics*, 5(2011):192, 2011.

Aad W Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

Aad W van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

Aad W. van der Vaart and Jon A. Wellner. *Weak Convergence and Empirical Processes*. Springer, New York, 1996.

Aad W Van Der Vaart and Jon A Wellner. Weak convergence. In *Weak convergence and empirical processes: with applications to statistics*, pages 16–28. Springer, 1996.

Stijn Vansteelandt and Oliver Dukes. Assumption-lean inference for generalised linear model parameters. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(3):657–685, 2022.

Stijn Vansteelandt and Els Goetghebeur. Causal inference with generalized structural mean models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 65(4):817–835, 2003.

Stijn Vansteelandt and Paweł Morzywołek. Orthogonal prediction of counterfactual outcomes. *arXiv preprint arXiv:2311.09423*, 2023.

Vladimir N Vapnik. An overview of statistical learning theory. *IEEE transactions on neural networks*, 10(5):988–999, 1999.

Vladimir N Vapnik and Y Alexey. Chervonenkis. on the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and its Applications*, 16(2):264–280, 1971.

Karel Vermeulen and Stijn Vansteelandt. Bias-reduced doubly robust estimation. *Journal of the American Statistical Association*, 110(511):1024–1036, 2015.

Karel Vermeulen and Stijn Vansteelandt. Data-adaptive bias-reduced doubly robust estimation. *The international journal of biostatistics*, 12(1):253–282, 2016.

R Von Mises. On the asymptotic distribution of differentiable statistical functions. *The Annals of Mathematical Statistics*, 18(3):309–348, 1947.

Vladimir Vovk and Ivan Petej. Venn-abers predictors. *arXiv preprint arXiv:1211.0025*, 2012.

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

Martin J Wainwright. *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge university press, 2019.

Lan Wang, Yu Zhou, Rui Song, and Ben Sherwood. Quantile-optimal treatment regimes. *Journal of the American Statistical Association*, 113(523):1243–1254, 2018.

Zeyi Wang, Wenxin Zhang, and Mark van der Laan. Super ensemble learning using the highly-adaptive-lasso. *arXiv preprint arXiv:2312.16953*, 2023.

Geoffrey S Watson. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 359–372, 1964.

Ted Westling, Peter Gilbert, and Marco Carone. Causal isotonic regression. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 82(3):719–747, 2020.

Ted Westling, Alex Luedtke, Peter B Gilbert, and Marco Carone. Inference for treatment-specific survival curves using machine learning. *Journal of the American Statistical Association*, 119(546):1541–1553, 2024.

Halbert White. Maximum likelihood estimation of misspecified models. *Econometrica: Journal of the econometric society*, pages 1–25, 1982.

Justin Whitehouse, Christopher Jung, Vasilis Syrgkanis, Bryan Wilder, and Zhiwei Steven Wu. Orthogonal causal calibration. *arXiv preprint arXiv:2406.01933*, 2024.

Nicholas T Williams, Oliver J Hines, and Kara E Rudolph. Riesz representers for the rest of us. *arXiv preprint arXiv:2507.19413*, 2025.

S. N. Wood. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, 73(1):3–36, 2011.

Simon N Wood. mgcv: Gams and generalized ridge regression for r. *R news*, 1(2): 20–25, 2001.

Marvin N. Wright and Andreas Ziegler. ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, 77 (1):1–17, 2017. doi: 10.18637/jss.v077.i01.

Yachong Yang, Arun Kumar Kuchibhotla, and Eric Tchetgen Tchetgen. Forster-warmuth counterfactual regression: A unified learning approach. *arXiv preprint arXiv:2307.16798*, 2023.

Yuhong Yang and Andrew Barron. Information-theoretic determination of minimax rates of convergence. *Ann. Statist.*, 27, 11 2000. doi: 10.1214/aos/1017939142.

Bianca Zadrozny and Charles Elkan. Obtaining calibrated probability estimates from decision trees and naive bayesian classifiers. In *Icml*, volume 1, pages 609–616, 2001.

Bianca Zadrozny and Charles Elkan. Transforming classifier scores into accurate multiclass probability estimates. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 694–699, 2002.

Jeffrey Zhang, Wei Li, Wang Miao, and Eric Tchetgen Tchetgen. Proximal causal inference without uniqueness assumptions. *Statistics & probability letters*, 198:109836, 2023.

Tianyu Zhang. *Sieve: Nonparametric Estimation by the Method of Sieves*, 2023. URL https://cran.r-project.org/web/packages/Sieve. R package, https://cran.r-project.org/web/packages/Sieve.

Tianyu Zhang and Noah Simon. Regression in tensor product spaces by the method of sieves. *Electronic Journal of Statistics*, 17(2):3660–3727, 2023.

Tong Zhang and Bin Yu. Boosting with early stopping: Convergence and consistency. *The Annals of Statistics*, 33(4):1538–1579, 2005. doi: 10.1214/009053605000000255.

Yuqian Zhang, Weijie Ji, and Jelena Bradic. Dynamic treatment effects: high-dimensional inference under model misspecification. *arXiv preprint arXiv:2111.06818*, 2021.

José Zubizarreta. Stable weights that balance covariates for estimation with incomplete outcome data. *Journal of the American Statistical Association*, 110:0–0, 04 2015. doi: 10.1080/01621459.2015.1023805.