

# A Convex Framework for Confounding Robust Inference

**Kei Ishikawa**

K.STONERIV@GMAIL.COM

*Department of Mathematics*

*ETH Zurich \**

*Rämistrasse 101, 8092 Zürich, Switzerland*

**Niao He**

NIAO.HE@INF.ETHZ.CH

*Department of Computer Science*

*ETH Zurich*

*Universitätstrasse 6, 8092 Zürich, Switzerland*

**Takafumi Kanamori**

KANAMORI@COMP.ISCT.AC.JP

*School of Computing*

*Institute of Science Tokyo*

*2-12-1 Okayama, Meguro-ku, Tokyo, Japan*

*RIKEN AIP*

*1-4-1 Nihonbashi, Chuo-ku, Tokyo, Japan*

**Editor:** Chris Wiggins

## Abstract

We study policy evaluation of offline contextual bandits subject to unobserved confounders. Sensitivity analysis methods are commonly used to estimate the policy value under the worst-case confounding scenario within a given uncertainty set. However, existing work often resorts to some coarse relaxation of the uncertainty set for the sake of tractability, leading to overly conservative estimation of the policy value. In this paper, we propose a general estimator that provides a sharp lower bound of the policy value using convex programming. The generality of our estimator enables various extensions such as sensitivity analysis using f-divergence, model selection with cross validation and information criterion, and robust policy learning with the sharp lower bound. Furthermore, our estimation method can be reformulated as an empirical risk minimization problem thanks to the strong duality, which enables us to provide strong theoretical guarantees of the proposed estimator using M-estimation techniques.

**Keywords:** convex optimization, causality, M-estimation, robust optimization, kernel methods.

## 1. Introduction

The offline contextual bandit, a simple but powerful model for decision-making, finds applications across a spectrum of domains from personalized medical treatment to recommendations and advertisements on online platforms. For the policy evaluation in this setting, the inverse probability weighting (IPW) method (Hirano and Imbens, 2001; Hirano et al., 2003) and its variants are commonly used. These methods rely on the so-called unconfoundedness

---

\*. Part of this work was done while the author was affiliated with Tokyo Institute of Technology (currently Science Tokyo).

assumption, presupposing full observability of pertinent variables to exclude any influence from unobserved factors on the action selection and the resulting reward (Rubin, 1974). However, in practice, this unconfoundedness assumption can easily be violated, since often times unobserved confounders are not recorded in the logged data. For instance, consider personalized medicine, where private information such as patients’ financial status is usually not recorded in medical records yet may impact both the choice of treatments and their efficacy. An expensive treatment option may be only available for wealthy patients, and higher financial status may also have a correlation with their health status. In such situations, if the treatment effect of the expensive option depends on the health status of the patients, we can potentially overestimate (or underestimate) the true treatment effect due to this hidden underlying correlation.

To address this, sensitivity analysis offers a solution by establishing a worst-case lower bound on the policy value. This involves defining an uncertainty set encompassing potential confounding situations and evaluating the policy value under the most adverse conditions within this set. Such a worst-case lower bound allows for robust decision-making in the presence of confounding. A wide range of sensitivity models has been studied over the years (Rosenbaum, 2002; Tan, 2006; Rosenbaum et al., 2010; Liu et al., 2013).

The main focus of our paper is the marginal sensitivity model by Tan (2006) and its extensions, which assumes similarity between the observational probability and the true probability of the action given the context. Recently, Zhao et al. (2019) introduced an elegant algorithm for Tan’s marginal sensitivity analysis using the linear fractional programming, which revitalized the study of this model. They also used the bootstrap method to provide an interpretable estimate of the lower bound. This approach was extended to policy learning in Kallus and Zhou (2018, 2021), where they established a statistical guarantee for confounding robust policy learning through lower bound maximization. Notably, their theoretical analysis included a uniform bound for the estimation error of a class of policies, unlike previous works that only focused on the evaluation of a single policy.

## 1.1 Motivation

These sensitivity analysis methods rely on algorithms using linear programming with relaxed constraints for their tractability, leading to overly conservative lower bound estimates of the policy value. Such estimates are not necessarily guaranteed to be a consistent estimator of the true lower bound. To obtain a sharp lower bound, conditional moment constraints must be leveraged, which are infinite dimensional linear constraints. Recently, Dorn and Guo (2022) analyzed these constraints and characterized the sharp lower bound of Tan’s marginal sensitivity model with a conditional quantile function of the reward distribution. Using this characterization, they proposed the first tractable algorithm to estimate the sharp lower bound that converges to the true lower bound of the policy value.

In this paper, we revisit the sharp estimation problem from a novel perspective. Instead of relying on conditional quantile functions, we propose a tractable approximation to the infinite dimensional conditional moment constraints, enabling the sharp lower bound estimation to be expressed directly as a convex optimization problem. Unlike the previous two-stage optimization method introduced by Dorn and Guo (2022), our formulation is more general and naturally includes their sharp estimator as a special case.

## 1.2 Contributions

The generality of our estimator expands the horizon of the sensitivity analysis, both in terms of theory and applications.

Our major theoretical contributions stem from the reformulation of the lower bound estimation as an empirical risk minimization using the strong duality. This enables us to leverage the standard proof techniques for M-estimation to derive consistency and asymptotics of our estimator, simplifying our proofs compared to the bespoke proofs in previous works (Kallus and Zhou, 2018, 2021). In particular, we are the first to prove a consistent policy learning guarantee for the sharp lower bound, which parallels the seminal result in Kallus and Zhou (2018, 2021) for the classic unsharp bound.

In terms of applications, our convex formulation is less restrictive compared to the two-stage approach of the previous sharp estimators (Dorn and Guo, 2022; Dorn et al., 2021) and offers possibilities for various extensions, such as generalization of the classic marginal sensitivity model (Tan, 2006) using f-divergence, policy learning with the sharp lower bound, and model selection using the cross validation or the generalized information criterion (Konishi and Kitagawa, 2008). Furthermore, our estimator can naturally handle both discrete and continuous treatment, setting it apart from conventional unsharp estimators (Tan, 2006; Zhao et al., 2019; Kallus and Zhou, 2018), which are unsuitable for cases involving continuous treatment.

Lastly, we highlight the difference between this paper from an earlier version by the authors (Ishikawa and He, 2023). All the analyses on asymptotic normality and its applications including hypothesis testing and model selection are the novel results of this paper. This paper also provides a more comprehensive study of the consistency and specification bias, by providing the consistent policy learning guarantee for a general Vapnik-Chervonenkis policy class with the sharp lower bound for the first time, and by characterizing the magnitude of the specification bias when the kernel cannot be chosen optimally to achieve zero bias.

## 1.3 Related works

The sensitivity analysis of average treatment effect using quantile regression by Dorn and Guo (2022); Dorn et al. (2021) was the first approach that consistently recover the true lower bound under Tan’s marginal sensitivity model, significantly sharpening bounds compared to traditional methods. Subsequently, this approach was also applied to the conditional average treatment effects (Oprescu et al., 2023) and individual treatment effect Jin et al. (2022). In parallel, confounding-robust policy evaluation techniques have evolved from evaluation to robust policy learning (Kallus and Zhou, 2018, 2021) and to sequential decision-making settings (Kallus and Zhou, 2020). Recently, the sharp bound introduced by Dorn and Guo (2022) has also been effectively adapted for sequential decision-making by Bruns-Smith and Zhou (2023); Bennett et al. (2024). However, extending this sharp bound approach specifically to policy learning has remained unexplored, and our paper is the first to address this gap.

In parallel, a growing literature has studied causal inference using modern machine learning methods, including instrumental variable estimation (Hartford et al., 2017), heterogeneous and individual treatment effect estimation (Wager and Athey, 2018; Shalit et al., 2017; Louizos et al., 2017), and debiased learning frameworks that combine machine learning

with classical causal inference (Chernozhukov et al., 2018). Kernel-based methods have also been widely used in causal inference, particularly for testing conditional independence and representing nonparametric conditional relationships (Gretton et al., 2005; Zhang et al., 2012; Peters et al., 2017).

Our convex reformulation of the infinite dimensional constraints is inspired by kernel methods (Schölkopf et al., 1997). The most relevant to our work is by Kremer et al. (2022), who used kernel methods to represent conditional moment restrictions and estimate model parameters. While their work focused on the dual formulation, we solve the primal but still exploit the dual in our theoretical analysis. Muandet et al. (2020a) also studied conditional moment restrictions, introducing a kernel-based test statistic structurally similar to ours but for hypothesis testing rather than uncertainty quantification. More broadly, the idea of using the kernel methods for constraints has been explored in areas such as fairness (Pérez-Suay et al., 2017), distributional robustness (Staib and Jegelka, 2019; Zhu et al., 2020), and shape-constrained regression (Aubin-Frankowski and Szabó, 2020). The kernel methods have also seen growing use in causal inference, including treatment effect estimation (Singh et al., 2024), instrumental variables (Singh et al., 2019; Muandet et al., 2020b; Zhang et al., 2023), negative controls (Singh, 2020; Kallus et al., 2021; Mastouri et al., 2021), and policy evaluation (Kallus, 2018).

#### 1.4 Organization of the paper

The remainder of the paper is organized as follows: In Section 2, we formally introduce the problem we study and propose a new sharp lower bound estimator of policy value. Section 3 studies the theoretical properties of the proposed estimator through reformulation of our estimation method via strong duality. In Section 4, we present our numerical experiments, and Section 5 concludes the paper.

## 2. Problem Settings and Proposed Method

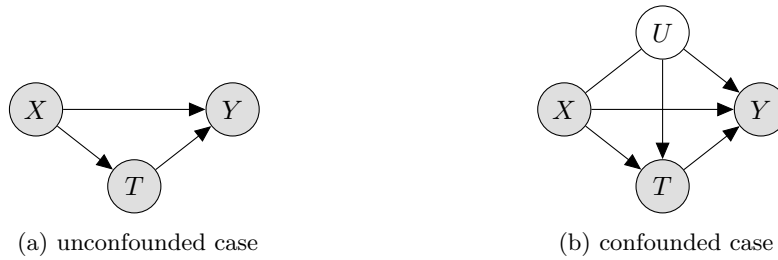
In this section, we start by introducing the formal definitions for confounded offline contextual bandits. Then, we discuss a sensitivity analysis model for its policy evaluation. Since the original formulation of this model requires an intractable optimization, we make a series of relaxations to motivate our proposed estimator.

Given the abundance of notations introduced herein, we offer a list of common notations used throughout the paper in Appendix A.

### 2.1 Confounded offline contextual bandits

Here, we formally define the confounded offline contextual bandits using the graphical models (Koller and Friedman, 2009; Pearl, 2009). Though a potential outcome framework (Rubin, 2005) is a popular alternative, we employed the graphical models as we find it easier to formulate our general modeling assumptions that accommodate both discrete or continuous action space.

The confounded offline contextual bandits are an extension of the standard offline contextual bandits that have an additional unobserved confounding variable. Let  $\mathcal{X}$  and  $\mathcal{T}$  be the context space and the action space. The model can be described as the probabilistic



**Figure 1:** Graphical model of unconfounded and confounded contextual bandit

model of the four variables: context  $X \in \mathcal{X}$ , confounder  $U \in \mathcal{U}$ , action (or treatment)  $T \in \mathcal{T}$ , and reward (or outcome)  $Y \in \mathbb{R}$ .<sup>1</sup> For some unknown base policy  $\pi_{\text{base}}(t|x, u)$ , we obtain an offline dataset of variables  $Y, T$ , and  $X$  from the following data generating process: Unlike the other variables confounder  $U$  is unobservable, and that introduces a confounding effect in the observable data. In confounded offline contextual bandits, the data are generated according to the following model:

$$\begin{aligned} X, U &\sim p(x, u), \\ T|X, U &\sim \pi_{\text{base}}(t|X, U), \\ Y|T, X, U &\sim p(y|T, X, U), \end{aligned} \tag{1}$$

but only  $Y, T$ , and  $X$  are observable and policy  $\pi_{\text{base}}(t|x, u)$  is unknown. This corresponds to the graphical model in Figure 1b, where the existence of  $U$  is the clear difference from the unconfounded graphical model in Figure 1a.<sup>2</sup>

To better understand variables, let us remember the earlier example of personalized medical treatment. The observational data  $\{(Y_i, T_i, X_i)\}_{i=1}^n$  corresponds to medical records. They include basic information about patients such as age, sex, and weight, which corresponds to context  $X$ . Additionally,  $T$  and  $Y$  are the medical treatment taken by the patients and its outcome. However, some information such as the financial status of the patients is not included in them. Yet, such information can influence the choice of treatment and the effectiveness of the treatment in either direct or indirect manners. Such a variable corresponds to confounder  $U$  in our model.

Now, we consider the offline evaluation of the policy from the observational data. Given new policy  $\pi(t|x)$ , we are interested in estimating the expectation of reward  $Y$  under the following probabilistic model:

$$\begin{aligned} X, U &\sim p(x, u), \\ T|X, U &\sim \pi(t|X), \\ Y|T, X, U &\sim p(y|T, X, U). \end{aligned} \tag{2}$$

Here, we only consider an observable policy  $\pi(t|x)$ , because it is trivially impossible to evaluate a policy that depends on unobserved variable  $U$  only using the offline data.

1. Though the outcome and the reward are not always the same in general, we assume them to be the same for simplicity.  
 2. Alternatively, we could consider the unconfounded case as the situation where confounder  $U$  is constant.

In confounded contextual bandits, offline evaluation of the above observable policy is still impossible without any further assumptions. To explain the reason, let's consider change of expectations

$$V(\pi) = \mathbb{E}_{T \sim \pi(\cdot|X)} [Y] = \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X,U)} \left[ \left( \frac{\pi(T|X)}{\pi_{\text{base}}(T|X,U)} \right) Y \right], \quad (3)$$

where  $\mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X,U)} [f(Y, T, X, U)]$  and  $\mathbb{E}_{T \sim \pi(\cdot|X)} [f(Y, T, X, U)]$  correspond to the expectations of  $f(Y, T, X, U)$  under the generative processes (1) and (2). Compared to the unconfounded case, the inverse probability in (3) has a dependence on unobserved variable  $U$ . This makes it impossible to construct a consistent estimator unlike the unconfounded case, where the inverse probability weighting estimator is consistent. As the observable variables are only  $Y$ ,  $T$ , and  $X$ , any valid estimator must depend only on them. Ideally, we would like to know conditional expectation of the inverse probability  $w(y, t, x) := \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X,U)} \left[ \frac{1}{\pi_{\text{base}}(T|X,U)} | Y = y, T = t, X = x \right]$ . Then, we can construct a valid unbiased and consistent estimator  $\hat{V}_w := \frac{1}{n} \sum_i w(Y_i, T_i, X_i) \pi(T_i|X_i) Y_i$ , because

$$V(\pi) = \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X,U)} \left[ \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X,U)} \left[ \frac{1}{\pi_{\text{base}}(T|X,U)} | Y, T, X \right] \pi(T|X) Y \right]. \quad (4)$$

However, in practice, it is impossible to know  $w(y, t, x)$  in the confounded offline contextual bandits, and it motivates the use of lower bound of the policy value.

Hereafter, for simplicity of notation and analysis, we assume that the conditional distribution of reward  $Y$  given  $T$ ,  $X$ , and  $U$  is continuous and that it yields a density function  $p(y|t, x, u)$  when such assumptions are appropriate. We will also assume that  $\pi(T|X, U) > 0$  almost surely under data generating process (1) so that the inverse probability weights are always well-defined. Additionally, we will use the following simplified notations. To indicate a part of model (1) that can be identified by the offline data, we use  $p_{\text{obs}}$  to indicate the observable distribution of (1) so that

$$p_{\text{obs}}(y, t, x) = \int p(y|t, x, u) \pi_{\text{base}}(t|x, u) p(x, u) du.^3$$

Similarly,  $p_{\text{obs}}(t|x)$  and  $p_{\text{obs}}(x)$  will denote the corresponding conditional and marginal distributions, and  $\mathbb{E}_{\text{obs}}[f(Y, T, X)] := \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X,U)} [f(Y, T, X)]$  represents the expectation of  $f(Y, T, X)$  with respect to  $p_{\text{obs}}(y, t, x)$ . To represent the empirical average that approximates  $\mathbb{E}_{\text{obs}}$ , we use  $\hat{\mathbb{E}}_n$  so that  $\hat{\mathbb{E}}_n[f(Y, T, X)] := \frac{1}{n} \sum_{i=1}^n f(Y_i, T_i, X_i)$  for any  $f(y, t, x)$ . Finally, throughout the paper, we assume that we know the true value of  $p_{\text{obs}}(t|x)$  in the estimation of the policy value and its lower (or upper) bounds for the sake of simplicity.<sup>4</sup>

3. We use notation  $\int$  to emphasize that unobserved confounder  $U$  could be both a summation over a discrete variable or an integral over a continuous variable. The same notation is used later for treatment  $T$ .

4. In the literature such as Kallus and Zhou (2018); Dorn and Guo (2022), it is commonly assumed that the probability of action given a context in observational data can be consistently estimated from observable data. This probability is often referred to as the "nominal propensity score." However, to maintain consistent notation, we will refer to it as the observational probability and denote it as  $p_{\text{obs}}(t|x)$  to highlight the tractable components of our formulas.

## 2.2 Uncertainty sets of base policies

A practical workaround to the above-mentioned issue is partial identification of the policy value under some reasonable assumption about confounding. More specifically, we first define uncertainty set  $\mathcal{E}$  of  $\pi_{\text{base}}(t|x, u)$  that we believe to contain all possible base policies underlying the observational data. Then we find infimum policy value  $V_{\text{inf}}$  (or supremum  $V_{\text{sup}}$ ) within the uncertainty set as

$$V_{\text{inf}}(\pi) := \inf_{\pi_{\text{base}} \in \mathcal{E}} \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X, U)} \left[ \left( \frac{\pi(T|X)}{\pi_{\text{base}}(T|X, U)} \right) Y \right].^5 \quad (5)$$

In the following, we introduce the conditional f-constraint, a class of constraints that can express a few reasonable classes of uncertainty sets.

Let  $f : \mathbb{R} \times \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}$  be a function satisfying that  $f_{t,x}(\cdot) := f(\cdot, t, x)$  is convex and  $f_{t,x}(1) = 0$  for any fixed  $t \in \mathcal{T}$  and  $x \in \mathcal{X}$ . Then, we define the uncertainty set of the conditional f-constraint<sup>6</sup> as

$$\mathcal{E}_{f_{t,x}} := \left\{ \begin{array}{l} \pi_{\text{base}}(t|x, u) : \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X, U)} \left[ f_{T,X} \left( \frac{p_{\text{obs}}(T|X)}{\pi_{\text{base}}(T|X, U)} \right) \right] \leq \gamma \\ \text{and} \\ \pi_{\text{base}}(t|x, u) \text{ is proper distribution} \end{array} \right\}. \quad (6)$$

Note that the lower bound for this uncertainty set is still intractable, as it involves the expectation with respect to unobserved variable  $U$ . This issue is addressed in the next subsection, where we introduce relaxation to these uncertainty sets to obtain a tractable formulation of the lower bound.

Though this class of uncertainty sets has not been studied in prior work, we focus on this type of uncertainty set because it is designed to include several practically relevant examples, such as box constraints and f-divergence constraints.

**Example 1 (Box constraints)** *The box constraints are the type of uncertainty set that has traditionally been adopted in the sensitivity analysis. They can be expressed as uncertainty set*

$$\mathcal{E}_{\text{box}} := \left\{ \begin{array}{l} \pi_{\text{base}}(t|x, u) : a_{\pi}(t, x) \leq \pi_{\text{base}}(t|x, u) \leq b_{\pi}(t, x) \\ \text{and} \\ \pi_{\text{base}}(\cdot|x, u) \text{ is proper distribution} \end{array} \right\} \quad (7)$$

for some  $a_{\pi}(t, x)$  and  $b_{\pi}(t, x)$ . This model includes the well-known marginal sensitivity model by Tan (2006) and most of its extensions (Zhao et al., 2019; Kallus and Zhou, 2018; Dorn and Guo, 2022). Tan considered a binary action space and assumed that the odds ratio of observational probability of action  $p_{\text{obs}}(t|x)$  and true base policy  $\pi_{\text{base}}(t|x, u)$  are not too far from 1 so that

$$\Gamma^{-1} \leq \frac{p_{\text{obs}}(t|x)(1 - \pi_{\text{base}}(t|x, u))}{(1 - p_{\text{obs}}(t|x))\pi_{\text{base}}(t|x, u)} \leq \Gamma.$$

5. Alternatively, one may take the infimum not only over  $\pi_{\text{base}}(t|x, u)$  but also over the joint distribution  $p(x, u)$  and the conditional outcome model  $p(y|t, x, u)$ . In contrast, our formulation assumes oracle access to the remaining components of the data-generating process. As we will show later, this assumption enables a sequence of relaxations that lead to a tractable lower-bound estimator.

6. Though this class of constraint contains the f-divergence constraint discussed later in Example 2, we do not call them "conditional f-divergence constraint". This is because it does not become f-divergence in general, for instance, in the case of the box constraints.

As  $p_{\text{obs}}(t|x)$  can be identified from the observational data, we can enforce such constraints by choosing  $a_\pi$  and  $b_\pi$  in (7) as

$$a_\pi(t, x) = \frac{1}{1 + \Gamma(1/p_{\text{obs}}(t|x) - 1)} \quad \text{and} \quad b_\pi(t, x) = \frac{1}{1 + \Gamma^{-1}(1/p_{\text{obs}}(t|x) - 1)}. \quad (8)$$

This model can be expressed in terms of the conditional  $f$ -constraint (6) by choosing

$$f_{t,x}(\tilde{w}) = I_{[a_{\tilde{w}}(t,x), b_{\tilde{w}}(t,x)]}(\tilde{w}) \quad (9)$$

for  $a_{\tilde{w}}(t, x) = p_{\text{obs}}(t|x)/b_\pi(t, x)$  and  $b_{\tilde{w}}(t, x) = p_{\text{obs}}(t|x)/a_\pi(t, x)$ ,  $\mathcal{E}_{f_{t,x}}$ , where  $I_{[a,b]}(\tilde{w})$  is a characteristic function<sup>7</sup> of interval  $[a, b]$  defined as

$$I_{[a,b]}(\tilde{w}) := \begin{cases} 0 & \text{if } \tilde{w} \in [a, b], \\ \infty & \text{otherwise.} \end{cases}$$

**Example 2 (f-divergence constraint)** Using the  $f$ -divergence, we can also define another class of uncertainty sets that belong to the family of conditional  $f$ -constraints. We call such a constraint and its corresponding sensitivity model the  $f$ -divergence constraint and the  $f$ -sensitivity model. To our knowledge, the majority of the existing literature only studies the box constraints, and we are the first to discuss the use of  $f$ -divergence constraint.<sup>8</sup>

Before introducing the new uncertainty set, we first recall the formal definition of  $f$ -divergence. Let  $p(t)$  and  $q(t)$  be the probability mass function over countable set  $\mathcal{T}$  (or the probability density function over  $\mathcal{T} \subseteq \mathbb{R}^d$  for some  $d$  with respect to the Lebesgue measure), and let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a convex function satisfying  $f(1) = 0$ . Then, the  $f$ -divergence between the distribution of  $p(t)$  and  $q(t)$  is defined as  $D_f[p||q] := \int_{\mathcal{T}} f\left(\frac{p(t)}{q(t)}\right) q(t) dt$ . The  $f$ -divergence is a rich class of divergence between probability distributions that includes many divergences such as the Kullback-Leibler (KL), squared Hellinger, and Pearson  $\chi^2$  divergences.

Using the  $f$ -divergence, we define a new class of uncertainty set as

$$\mathcal{E}_f := \left\{ \begin{array}{l} \mathbb{E}_{X,U} [D_f[p_{\text{obs}}(t|X)||\pi_{\text{base}}(t|X,U)]] \leq \gamma \\ \pi_{\text{base}}(t|x, u) : \quad \text{and} \\ \pi_{\text{base}}(t|x, u) \text{ is proper distribution} \end{array} \right\}, \quad (10)$$

where the expectation  $\mathbb{E}_{X,U}$  can be taken with respect to  $p(x, u)$  in both (1) and (2), regardless of the policy. This formulation naturally fits into the conditional  $f$ -constraints (6) as the divergence term becomes a simple expectation as  $\mathbb{E}_{X,U} [D_f[p_{\text{obs}}(t|X)||\pi_{\text{base}}(t|X,U)]] = \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X,U)} \left[ f\left(\frac{p_{\text{obs}}(T|X)}{\pi_{\text{base}}(T|X,U)}\right) \right]$ .

The key distinction between  $f$ -divergence constraints and box constraints lies in their non-uniform nature. While box constraints impose uniform bounds on the odds ratio  $\frac{p_{\text{obs}}(t|x)}{\pi_{\text{base}}(t|x,u)}$  with respect to context  $X$ ,  $f$ -divergence constraints only limit the overall deviation of the policy in aggregate. As a result, they allow for base policies that can vary significantly for a limited range of context values  $X$ . Practically, box constraints are appropriate when one wants a worst-case, uniform robustness guarantee across all contexts. In contrast,  $f$ -divergence constraints are useful when one expects confounding to be concentrated on a subset of the population and prefers to control the average deviation globally.

7. Please be aware that we use  $\mathbb{1}_{[a,b]}$  to represent indicator function that takes value in  $\{0, 1\}$ .

8. Indeed, Jin et al. (2022) considered a similar use of  $f$ -divergence for sensitivity analysis. However, their model makes a different type of confounding assumption, as discussed in Appendix B.

### 2.3 Relaxation of the uncertainty sets by reparametrization

Having defined the uncertainty set of our interest, we now consider the computation method of lower bound  $V_{\text{inf}}$ . Unfortunately, directly solving (5) is often too difficult in practice, so we instead solve a relaxed problem that is more tractable. First, we introduce reparametrization  $w(y, t, x) = \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X, U)} \left[ \frac{1}{\pi_{\text{base}}(T|X, U)} | Y = y, T = t, X = x \right]$ . By relaxing the uncertainty set of unknown base policy  $\pi_{\text{base}}(t|x, u)$ , we can construct an uncertainty set of  $w(y, t, x)$  and obtain a convex relaxation of original problem (5).

Without the relaxation, the uncertainty set of base policy  $\mathcal{E}_{f_{t,x}}$  can be translated to the uncertainty set of the reparametrized weights as

$$\mathcal{W}_{f_{t,x}} := \left\{ w(y, t, x) : \begin{array}{l} w(y, t, x) = \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X, U)} \left[ \frac{1}{\pi_{\text{base}}(T|X, U)} | Y = y, T = t, X = x \right] \\ \text{and} \\ \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X, U)} \left[ f_{T, X} \left( \frac{p_{\text{obs}}(T|X)}{\pi_{\text{base}}(T|X, U)} \right) \right] \leq \gamma \\ \text{and} \\ \pi_{\text{base}}(t|x, u) \text{ is proper distribution} \end{array} \right\}. \quad (11)$$

Using (4), we can obtain the lower bound as

$$V_{\text{inf}}(\pi) = \inf_{w \in \mathcal{W}_{f_{t,x}}} \mathbb{E}_{\text{obs}} [w(Y, T, X) \pi(T|X) Y].$$

This uncertainty set requires that the conditional f-constraint is satisfied and that the base policy  $\pi_{\text{base}}(t|x, u)$  is proper distribution. However, in practice, both conditions are intractable. Therefore, we will consider the relaxation of these conditions.

#### RELAXATION OF THE CONDITIONAL F-CONSTRAINT

Let us consider the first condition of (11) on conditional f-divergence. With Jensen's inequality, we have

$$\mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X, U)} \left[ f_{T, X} \left( \frac{p_{\text{obs}}(T|X)}{\pi_{\text{base}}(T|X, U)} \right) \right] \geq \mathbb{E}_{\text{obs}} \left[ f_{T, X} \left( \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X, U)} \left[ \frac{p_{\text{obs}}(T|X)}{\pi_{\text{base}}(T|X, U)} | Y, T, X \right] \right) \right],$$

where the equality holds when there's  $\pi_{\text{base}}(t|x, u) = p_{\text{obs}}(t|x)$  (i.e., no confounding) or when  $f_{T, X}(\cdot)$  is linear almost surely.<sup>9</sup> Thus, we relax condition  $\mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X, U)} \left[ f_{T, X} \left( \frac{p_{\text{obs}}(T|X)}{\pi_{\text{base}}(T|X, U)} \right) \right] \leq \gamma$  in (11) to

$$\mathbb{E}_{\text{obs}} [f_{T, X} (p_{\text{obs}}(T|X) w(Y, T, X))] \leq \gamma. \quad (12)$$

Here, we must emphasize that the above constraint is convex with respect to  $w(y, t, x)$ . This convexity translates to the empirical version of the above constraint, and enables us to estimate the lower bound by solving convex programming.

9. Indeed, the box constraints satisfy this linear condition and no relaxation is made. On the other hand, the f-divergence constraint relies on strictly convex functions, and thus the equality of Jensen's inequality does not hold.

Additionally, when  $f_{t,x}$  does not depend on  $t$  and  $x$ , there is a simple trick to ensure that the left hand side of (12) is non-negative. By imposing  $\mathbb{E}_{\text{obs}}[p_{\text{obs}}(T|X)w(Y, T, X)] = 1$ , we can guarantee  $0 = f(1) = f(\mathbb{E}_{\text{obs}}[p_{\text{obs}}(T|X)w(Y, T, X)]) \leq \mathbb{E}_{\text{obs}}[f(p_{\text{obs}}(T|X)w(Y, T, X))]$  due to Jensen's inequality. In the numerical experiments, we observed that this helps to obtain a more stable implementation of the f-sensitivity model.

#### RELAXATION OF THE DISTRIBUTIONAL CONSTRAINTS

Now we consider the relaxation of the second condition in (11) that  $\pi_{\text{base}}(t|x, u)$  is proper distribution, which is equivalent to

$$\int_{\mathcal{T}} \pi_{\text{base}}(t'|x, u) dt' = 1 \text{ and } \pi_{\text{base}}(t|x, u) \geq 0 \text{ for any } t \in \mathcal{T}, x \in \mathcal{X}, \text{ and } u \in \mathcal{U}. \quad (13)$$

We relax constraint (13) as conditional moment constraints

$$\mathbb{E}_{\text{obs}}[w(Y, T, X)|T = t, X = x] \cdot p_{\text{obs}}(t|x) = 1 \text{ for any } t \in \mathcal{T} \text{ and } x \in \mathcal{X} \quad (14)$$

and non-negativity condition

$$w(y, t, x) \geq 0 \text{ for any } y \in \mathcal{Y}, t \in \mathcal{T}, \text{ and } x \in \mathcal{X}.^{10}$$

To simplify the treatment of the non-negativity condition, we hereafter assume that the non-negativity condition is already imposed by conditional f-constraint (12) so that  $f_{t,x}(u) = \infty$  for any  $u < 0$ . This can be assumed without loss of generality because for any conditional-f constraint  $\mathbb{E}_{\text{obs}}[f_{t,x}(w(Y, T, X)p_{\text{obs}}(T|X))] \leq \gamma$ , we can consider a new f-constraint  $\mathbb{E}_{\text{obs}}[\tilde{f}_{t,x}(w(Y, T, X)p_{\text{obs}}(T|X))] \leq \gamma$  for  $\tilde{f}_{t,x}(v) = f_{t,x}(v) + I_{[0, \infty)}(v)$ .

The validity of this relaxation can be checked as

$$\begin{aligned} & \mathbb{E}_{\text{obs}}[w(Y, T, X)|T = t, X = x] \cdot p_{\text{obs}}(t, x) \\ &= \mathbb{E}_{\text{obs}} \left[ \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X, U)} \left[ \frac{1}{\pi_{\text{base}}(T|X, U)} |Y, T, X \right] |T = t, X = x \right] \cdot p_{\text{obs}}(t, x) \\ &= \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X, U)} \left[ \frac{1}{\pi_{\text{base}}(T|X, U)} |T = t, X = x \right] \cdot p_{\text{obs}}(t, x) \\ &= \int_{\mathcal{U}} \frac{1}{\pi_{\text{base}}(t|x, u)} p_{\pi_{\text{base}}}(u|t, x) du \cdot p_{\text{obs}}(t, x) \\ &= \int_{\mathcal{U}} \frac{1}{\pi_{\text{base}}(t|x, u)} p_{\pi_{\text{base}}}(t, x, u) du \\ &= \int_{\mathcal{U}} p(x, u) du = p_{\text{obs}}(x) \end{aligned}$$

and the fact that  $p_{\text{obs}}(t|x) = p_{\text{obs}}(t, x)/p_{\text{obs}}(x)$ . Here,  $p_{\pi_{\text{base}}}(t, x, u)$  and  $p_{\pi_{\text{base}}}(u|t, x)$  denote the joint and conditional distribution of  $T$ ,  $X$ , and  $U$  under confounded contextual bandits

10. When  $\mathcal{T}$  is discrete, it is possible to further tighten this condition as  $w \geq 1$  by using the fact that  $\pi_{\text{base}}(t|x, u) \leq 1$ . However, we leave this condition as it is so that it can handle cases where  $\mathcal{T}$  is not discrete.

model (1). To summarize, we combine these conditional moment constraints (CMC) with relaxed conditional f-constraint (12) to define relaxed uncertainty set

$$\mathcal{W}_{f_t,x}^{\text{CMC}} := \left\{ w(y, t, x) : \begin{array}{l} \mathbb{E}_{\text{obs}} [f_{T,X} (p_{\text{obs}}(T|X)w(Y, T, X))] \leq \gamma \\ \text{and} \\ \mathbb{E}_{\text{obs}} [w(Y, T, X) | T = t, X = x] \cdot p_{\text{obs}}(t|x) = 1 \text{ for any } t \in \mathcal{T} \text{ and } x \in \mathcal{X} \end{array} \right\}$$

and its corresponding lower bound

$$V_{\text{inf}}^{\text{CMC}}(\pi) := \inf_{w \in \mathcal{W}_{f_t,x}^{\text{CMC}}} \mathbb{E}_{\text{obs}} [w(Y, T, X)\pi(T|X)Y]. \quad (15)$$

**Remark 1 (A comparison to ZSB sensitivity model)** *Until the recent work by Dorn and Guo (2022), unsharp estimators of the lower bound had been the de facto standard in sensitivity analysis. Rather than adopting the conditional moment constraint (14) as relaxation, earlier works — including Tan (2006); Zhao et al. (2019); Kallus and Zhou (2018, 2021) — relied on the following looser relaxations, leading to unsharp estimates.*

$$\mathbb{E}_{\text{obs}}[\mathbb{1}_{T=t}w(Y, T, X)] = 1 \text{ for any } t \in \mathcal{T},$$

where  $\mathbb{1}$  denotes the indicator function for event  $A$ . This condition can be obtained from conditional moment constraints (14) as  $\mathbb{E}_{\text{obs}}[\mathbb{1}_{T=t}w(Y, T, X)] = \mathbb{E}_{\text{obs}}[\mathbb{1}_{T=t}\mathbb{E}_{\text{obs}}[w(Y, T, X)|T, X]] = \mathbb{E}_{\text{obs}}[\mathbb{1}_{T=t}/p_{\text{obs}}(T|X)] = 1$  for any  $t \in \mathcal{T}$ . Following the naming convention in Dorn and Guo (2022), we call these constraints the ZSB constraints after the authors of Zhao et al. (2019).

Combining the above with the relaxation of the conditional f-constraint as in (12), the following uncertainty set with the ZSB constraints becomes

$$\mathcal{W}_{f_t,x}^{\text{ZSB}} := \left\{ w(y, t, x) : \begin{array}{l} \mathbb{E}_{\text{obs}} [f_{T,X} (p_{\text{obs}}(T|X)w(Y, T, X))] \leq \gamma \\ \text{and} \\ \mathbb{E}_{\text{obs}}[\mathbb{1}_{T=t}w(Y, T, X)] = 1 \text{ for any } t \in \mathcal{T} \end{array} \right\},$$

with associated lower bound

$$V_{\text{inf}}^{\text{ZSB}}(\pi) := \inf_{w \in \mathcal{W}_{f_t,x}^{\text{ZSB}}} \mathbb{E}_{\text{obs}} [w(Y, T, X)\pi(T|X)Y]. \quad (16)$$

In practice, for the ZSB constraint to be tractable, we need to assume that action space  $\mathcal{T}$  is discrete and finite. This is in contrast to our approach, which directly approximates the conditional moment constraints and can handle both discrete and continuous action spaces.

Naturally, for the uncertainty sets discussed, one can show inclusion relations  $\mathcal{W}_{f_t,x} \subseteq \mathcal{W}_{f_t,x}^{\text{CMC}} \subseteq \mathcal{W}_{f_t,x}^{\text{ZSB}}$ . The former inclusion follows from the definition of  $\mathcal{W}_{f_t,x}^{\text{CMC}}$ , and the latter follows from the fact that the ZSB constraint follows from the conditional moment constraints as discussed above. Under the assumption that the original sensitivity model is correctly specified — meaning the true base policy  $\pi_{\text{base}}(t|x, u)$  lies within the original uncertainty set  $\mathcal{E}_{f_t,x}$  — we can ensure that the tightness of the resulting lower bounds aligns with the order of relaxation, i.e.,  $V \geq V_{\text{inf}} \geq V_{\text{inf}}^{\text{CMC}} \geq V_{\text{inf}}^{\text{ZSB}}$ .

11

11. Here, it is worth considering under what kind of setting the second equality  $V_{\text{inf}} = V_{\text{inf}}^{\text{CMC}}$  holds so that the lower bound of the conditional moment constraints is tight. Surprisingly, a recent work by Dorn

## 2.4 Low-rank approximation of the conditional moment constraints

Now, we introduce an empirical approximation of conditional moment constraints (14) using low-rank approximation.

Let us introduce another re-parametrization  $e(y, t, x) := p_{\text{obs}}(t|x)w(y, t, x) - 1$  for further tractability so that conditional moment constraints (14) can be written as

$$\mathbb{E}_{\text{obs}}[e(Y, T, X)|T = t, X = x] = 0 \quad \text{for any } t \in \mathcal{T} \text{ and } x \in \mathcal{X}.$$

This is equivalent to orthogonality condition  $\mathbb{E}_{\text{obs}}[e(Y, T, X)\psi(T, X)] = 0$  for any function  $\psi(t, x)$ . However, imposing this orthogonality condition for all possible  $\psi$  is infeasible if we impose the empirical version of the orthogonality constraints, as there are infinite choices of  $\psi$ . To cope with this issue, we apply low-rank approximation to the orthogonality conditions above. With a collection of basis functions  $\{\psi_d(t, x)\}_{d=1}^D$ , we define the kernel conditional moment constraints (KCMC) for the population as

$$\text{KCMC} \stackrel{\text{def}}{\Leftrightarrow} \mathbb{E}_{\text{obs}}[e(Y, T, X)\psi_d(T, X)] = 0 \quad \text{for any } d = 1, \dots, D. \quad (17)$$

We call this approximation the kernel conditional moment constraints because it can be motivated from the existing kernel methods. In the preliminary version of our work (Ishikawa and He, 2023), we show that the low-rank kernel ridge regression can also be used to motivate the KCMC. For example, it could be considered a version of the kernel formulation of conditional moment constraints by Kremer et al. (2022) and Muandet et al. (2020a) using low-rank kernel  $k((t, x), (t', x')) := \sum_{d=1, \dots, D} \psi_d(t_1, x_1)\psi_d(t_2, x_2)$ .<sup>12</sup> Additionally, as we discuss later in Example 3, the principal components of the kernel PCA (Schölkopf et al., 1997) can be used as a reasonable collection of basis functions. When we take first  $D$  principal components found by kernel PCA as the basis functions for KCMC, we can prove that KCMC can approximate the original conditional moment constraints at an arbitrary precision as  $D$  and  $n$  go to infinity.

Analogously, the empirical version of the KCMC is also defined by replacing  $\mathbb{E}_{\text{obs}}$  with  $\hat{\mathbb{E}}_n$ . Finally, combining the KCMC approximation (17) and relaxed conditional f-constraint (12), we can define the uncertainty set

$$\mathcal{W}_{f_{t,x}}^{\text{KCMC}} = \left\{ w(y, t, x) : \begin{array}{l} \mathbb{E}_{\text{obs}}[f_{T,X}(w(Y, T, X)p_{\text{obs}}(T|X))] \leq \gamma \\ \text{and} \\ \mathbb{E}_{\text{obs}}[(w(Y, T, X)p_{\text{obs}}(T|X) - 1)\psi_d] = 0 \quad \text{for any } d = 1, \dots, D. \end{array} \right\}$$

---

and Guo (2022, Section 3.2) showed that this equality holds for average treatment effect estimation with Tan's marginal sensitivity model (Tan, 2006). They showed that there exists minimizer  $w^*(y, t, x)$  of  $\text{ATE}(w) := \mathbb{E}_{\text{obs}}[w(Y, T, X)\mathbb{1}_{T=1}Y] - \mathbb{E}_{\text{obs}}[w(Y, T, X)\mathbb{1}_{T=0}Y]$ , which is realizable so that there exists  $p(y|t, x, u)$ ,  $\pi_{\text{base}}(t|x, u)$ , and  $p(x, u)$  that is compatible with any  $p_{\text{obs}}(y, t, x)$  and minimizer  $w^*(y, t, x)$ . However, they did not provide any realizability results for more general settings discussed in this paper, such as the evaluation of general policy with non-binary action space and conditional f-constraint. This is an open question we could not address in our paper.

12. Instead of the low-rank approximation, we could use the maximum moment restriction (MMR) as in Kremer et al. (2022) and Muandet et al. (2020a). The MMR is a quadratic form that quantifies the violation of the conditional moment and upper bounding it leads to an alternative convex uncertainty set. However, in terms of the theoretical analysis, their approach is more technically involved, and deriving asymptotic results comparable to ours becomes substantially more complicated than in our finite-dimensional, low-rank setting, where our theoretical results can be derived as special cases of classic asymptotics in a finite-dimensional parameter space.

and the associated lower bound with the KCMC as

$$V_{\inf}^{\text{KCMC}}(\pi) = \inf_{w \in \mathcal{W}_{f_t, x}^{\text{KCMC}}} \mathbb{E}_{\text{obs}}[w(Y, T, X)\pi(T|X)Y].$$

Likewise, we use its empirical version and define the KCMC estimator as

$$\hat{V}_{\inf}^{\text{KCMC}}(\pi) = \inf_{w \in \hat{\mathcal{W}}_{f_t, x}^{\text{KCMC}}} \hat{\mathbb{E}}_n[w(Y, T, X)\pi(T|X)Y], \quad (18)$$

where

$$\hat{\mathcal{W}}_{f_t, x}^{\text{KCMC}} = \left\{ w(y, t, x) : \begin{array}{l} \hat{\mathbb{E}}_n[f_{T, X}(w(Y, T, X)p_{\text{obs}}(T|X))] \leq \gamma \\ \text{and} \\ \hat{\mathbb{E}}_n[(w(Y, T, X)p_{\text{obs}}(T|X) - 1)\psi_d] = 0 \text{ for any } d = 1, \dots, D \end{array} \right\}.$$

As can be seen, the KCMC estimator is the solution of a convex programming with a tractable feasibility set. Therefore, our estimator can easily implemented using any convex optimization solvers such as CVXPY (Diamond and Boyd, 2016).

### 3. Theoretical Analysis

In this section, we study the property of lower bound estimate of the policy value with the KCMC estimator. We begin the analysis of the KCMC estimator by deriving the convex dual of the above problems and characterizing their minimizers. Then, we study the specification error of the KCMC lower bound  $|V_{\inf}^{\text{CMC}} - V_{\inf}^{\text{KCMC}}|$ . We provide a condition on the orthogonal function set  $\{\psi_d(t, x)\}_{d=1}^D$  under which the specification error can be bounded. Next, we show the consistency of the KCMC estimator in policy evaluation by showing the equivalence of the KCMC estimator with a standard M-estimator. We further prove consistency guarantees for policy learning with the KCMC estimator, in the cases of a concave policy class and the Vapnik-Chevonenkis policy class. Lastly, we derive the asymptotic normality of the KCMC estimator and discuss its application to the construction of a confidence interval and model selection. For the sake of conciseness, we defer some proofs and detailed regularity conditions to Appendix C.

Before further discussion, we introduce several simplifications of the notations. We omit subscripts of  $\mathcal{W}_{f_t, x}^{\text{CMC}}$ ,  $\mathcal{W}_{f_t, x}^{\text{KCMC}}$ ,  $\hat{\mathcal{W}}_{f_t, x}^{\text{KCMC}}$  and  $\mathbb{E}_{\text{obs}}$ , unless they are unclear from the context. We also introduce  $r(y, t, x) = \left(\frac{\pi(t|x)}{p_{\text{obs}}(t|x)}\right) \cdot y$  and reparametrization  $\tilde{w}(y, t, x) = p_{\text{obs}}(t|x)w(y, t, x)$ .

Furthermore, we briefly review the subgradient and the Fenchel conjugate here, as we will make heavy use of them in this section. The subgradient of convex function  $f$  is represented by  $\partial f$ . When we apply the addition operator to the subgradient, it represents the Minkowski sum. Other operations to the subgradient such as multiplication are similarly defined. The Fenchel conjugate of  $f : \mathbb{R} \rightarrow \mathbb{R}$  is defined as  $f^*(v) := \sup_u \{uv - f(u)\}$ , and it has a few important properties. The Fenchel conjugate is always convex because the supremum of a family of convex functions is convex.<sup>13</sup> Additionally, there exists maximizer  $u^*$  that solves

13. In this case,  $v \mapsto vu - f(u)$  is linear, and therefore, is convex.

$f^*(v) = \max_u \{vu - f(u)\}$  and it satisfies

$$u^* \in \partial f^*(v) \quad (19)$$

if  $f$  is closed and convex. Function  $f$  is closed and convex if its epigraph  $\text{epi}(f) := \{(u, t) : u \in \text{dom}(f), f(u) \leq t\}$  is closed and convex, and these conditions are satisfied in our problem. A more thorough treatment of the subgradient and the Fenchel conjugate can be found in (Boyd et al., 2004).

### 3.1 Characterization of the solution

In this subsection, we derive explicit formulae for the minimizers that give three lower bounds  $V_{\text{inf}}^{\text{CMC}}$ ,  $V_{\text{inf}}^{\text{KCMC}}$ , and  $\hat{V}_{\text{inf}}^{\text{KCMC}}$ , which are

$$\begin{aligned} w_{\text{CMC}}^* &= \arg \min_{w \in \mathcal{W}^{\text{CMC}}} \mathbb{E}[w(Y, T, X)\pi(T|X)Y], \\ w_{\text{KCMC}}^* &= \arg \min_{w \in \mathcal{W}^{\text{KCMC}}} \mathbb{E}[w(Y, T, X)\pi(T|X)Y], \\ \hat{w}_{\text{KCMC}} &= \arg \min_{w \in \hat{\mathcal{W}}^{\text{KCMC}}} \hat{\mathbb{E}}_n[w(Y, T, X)\pi(T|X)Y]. \end{aligned} \quad (20)$$

Here, we know that these problems have minimizers because the above problems are minimizations of linear objectives under convex constraints. Furthermore, we know that the strong duality holds for the above convex optimizations, as their feasibility sets have a non-empty relative interior, satisfying Slater's constraint qualification. Using these properties, we obtain the following lemma:

**Lemma 1 (Characterization of solutions)** *Let  $w_{\text{CMC}}^*$ ,  $w_{\text{KCMC}}^*$ , and  $\hat{w}_{\text{KCMC}}$  be defined as in (20). Then, there exist function  $\eta_{\text{CMC}} : \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}$ , vectors  $\eta_{\text{KCMC}}, \eta'_{\text{KCMC}} \in \mathbb{R}^D$ , and constants  $\eta_f, \eta'_f, \eta''_f > 0$  such that*

$$w_{\text{CMC}}^*(y, t, x) \in \left( \frac{1}{p_{\text{obs}}(t|x)} \right) \partial f_{t,x}^* \left( \frac{\eta_{\text{CMC}}(t, x) - r(y, t, x)}{\eta_f} \right), \quad (21)$$

$$w_{\text{KCMC}}^*(y, t, x) \in \left( \frac{1}{p_{\text{obs}}(t|x)} \right) \partial f_{t,x}^* \left( \frac{\eta_{\text{KCMC}}^T \psi(t, x) - r(y, t, x)}{\eta'_f} \right), \quad (22)$$

$$\hat{w}_{\text{KCMC}}(y, t, x) \in \left( \frac{1}{p_{\text{obs}}(t|x)} \right) \partial f_{t,x}^* \left( \frac{\eta'_{\text{KCMC}}^T \psi(t, x) - r(y, t, x)}{\eta''_f} \right). \quad (23)$$

**Proof** See below. ■

#### CHARACTERIZATION OF $w_{\text{CMC}}^*$

By using the strong duality, we can transform the original problem for  $V_{\text{inf}}^{\text{CMC}}$  as

$$V_{\text{inf}}^{\text{CMC}} = \inf_{w \in \mathcal{W}^{\text{CMC}}} \mathbb{E}[\tilde{w}(Y, T, X)r(Y, T, X)]$$

$$\begin{aligned}
 &= \inf_{\tilde{w}} \sup_{\substack{\eta_{\text{CMC}}: \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}, \\ \eta_f \geq 0}} \mathbb{E}[\tilde{w}r] - \mathbb{E}[\eta_{\text{CMC}}(T, X)(\tilde{w} - 1)] + \eta_f (\mathbb{E}[f_{T, X}(\tilde{w})] - \gamma) \\
 &= \sup_{\substack{\eta_{\text{CMC}}: \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}, \\ \eta_f \geq 0}} \inf_{\tilde{w}} \mathbb{E}[\tilde{w}r] - \mathbb{E}[\eta_{\text{CMC}}(T, X)(\tilde{w} - 1)] + \eta_f (\mathbb{E}[f_{T, X}(\tilde{w})] - \gamma) \\
 &= \sup_{\substack{\eta_{\text{CMC}}: \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}, \\ \eta_f > 0}} -\eta_f \gamma + \mathbb{E}[\eta_{\text{CMC}}] - \eta_f \mathbb{E} \left[ \sup_{\tilde{w}} \left\{ \left( \frac{\eta_{\text{CMC}} - r}{\eta_f} \right) \tilde{w} - f_{T, X}(\tilde{w}) \right\} \right] \\
 &= \sup_{\substack{\eta_{\text{CMC}}: \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}, \\ \eta_f > 0}} -\eta_f \gamma + \mathbb{E}[\eta_{\text{CMC}}] - \eta_f \mathbb{E} \left[ f_{T, X}^* \left( \frac{\eta_{\text{CMC}} - r}{\eta_f} \right) \right], \tag{24}
 \end{aligned}$$

assuming that optimal  $\eta_f$  is positive. This assumption does not hold only when conditional f-constraint (12) is not tight. However, in such cases,

$$\begin{aligned}
 V_{\text{inf}}^{\text{CMC}} &= \sup_{\eta_{\text{CMC}}: \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}} \inf_{\tilde{w}} \mathbb{E}[\tilde{w}r] - \mathbb{E}[\eta_{\text{CMC}}(T, X)(\tilde{w} - 1)] \\
 &= \sup_{\eta_{\text{CMC}}: \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}} \inf_{\tilde{w}} \mathbb{E}[\tilde{w}(Y, T, X)(r(Y, T, X) - \eta_{\text{CMC}}(T, X))] + \mathbb{E}[\eta_{\text{CMC}}(T, X)] \tag{25}
 \end{aligned}$$

and in order for the supremum to be reached for  $\eta_f = 0$ , we need  $r(Y, T, X) - \eta_{\text{CMC}}(T, X) = 0$  almost surely, which is not satisfied in practice. Therefore, for simplicity in our proofs, we will focus on cases where  $\eta_f > 0$ .

Now, as primal solution  $w_{\text{CMC}}^*$  must satisfy the Karush-Kuhn-Tucker (KKT) conditions, we can take the maximizers of (24) as  $\eta_f^*$  and  $\eta_{\text{CMC}}^*(t, x)$ . Using property (19) of the Fenchel conjugate, we can obtain solution form of  $w_{\text{CMC}}^*$

$$w_{\text{CMC}}^*(y, t, x) \in \left( \frac{1}{p_{\text{obs}}(t|x)} \right) \partial f_{t, x}^* \left( \frac{\eta_{\text{CMC}}^*(t, x) - r(y, t, x)}{\eta_f^*} \right),$$

which proves (21). Here, factor  $1/p_{\text{obs}}(t|x)$  in front of the subgradient appears because of reparametrization  $\tilde{w}(y, t, x) = p_{\text{obs}}(t|x)w(y, t, x)$ .

#### CHARACTERIZATION OF $w_{\text{KCMC}}^*$

Now, we derive the characterization of  $w_{\text{KCMC}}^*$ . Let  $\boldsymbol{\psi}(t, x) = (\psi_1(T, X), \dots, \psi_D(T, X))^T$ . We can use exactly the same technique as the above and reach a similar characterization of the solution as

$$\begin{aligned}
 V_{\text{inf}}^{\text{KCMC}} &= \inf_{w \in \mathcal{W}^{\text{KCMC}}} \mathbb{E}[\tilde{w}(Y, T, X)r(Y, T, X)] \\
 &= \inf_{\tilde{w}} \sup_{\substack{\eta_{\text{KCMC}} \in \mathbb{R}^D, \\ \eta_f \geq 0}} \mathbb{E}[\tilde{w}r] - \mathbb{E}[(\tilde{w} - 1)\eta_{\text{KCMC}}^T \boldsymbol{\psi}] + \eta_f (\mathbb{E}[f_{T, X}(\tilde{w})] - \gamma) \\
 &= \sup_{\substack{\eta_{\text{KCMC}} \in \mathbb{R}^D, \\ \eta_f \geq 0}} \inf_{\tilde{w}} \mathbb{E}[\tilde{w}r] - \mathbb{E}[(\tilde{w} - 1)\eta_{\text{KCMC}}^T \boldsymbol{\psi}] + \eta_f (\mathbb{E}[f_{T, X}(\tilde{w})] - \gamma) \\
 &= \sup_{\substack{\eta_{\text{KCMC}} \in \mathbb{R}^D, \\ \eta_f > 0}} -\eta_f \gamma + \eta_{\text{KCMC}}^T \mathbb{E}[\boldsymbol{\psi}] - \eta_f \mathbb{E} \left[ \sup_{\tilde{w}} \left\{ \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r}{\eta_f} \right) \tilde{w} - f_{T, X}(\tilde{w}) \right\} \right]
 \end{aligned}$$

$$= \sup_{\substack{\eta_{\text{KCMC}} \in \mathbb{R}^D, \\ \eta_f > 0}} -\eta_f \gamma + \eta_{\text{KCMC}}^T \mathbb{E}[\boldsymbol{\psi}] - \eta_f \mathbb{E} \left[ f_{T,X}^* \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r}{\eta_f} \right) \right]. \quad (26)$$

Now, using the maximizers of dual problem (26)  $\eta_f^*$  and  $\eta_{\text{KCMC}}^*$ , we can obtain a characterization of  $w_{\text{KCMC}}^*$  as

$$w_{\text{KCMC}}^*(y, t, x) \in \left( \frac{1}{p_{\text{obs}}(t|x)} \right) \partial f^* \left( \frac{\eta_{\text{KCMC}}^{*T} \boldsymbol{\psi}(t, x) - r(y, t, x)}{\eta_f^*} \right),$$

which proves (22).

#### CHARACTERIZATION OF $\hat{w}_{\text{KCMC}}$

Again, using the same techniques, we can derive the characterization of  $\hat{w}_{\text{KCMC}}$ . By exchanging  $\mathbb{E}$  with  $\hat{\mathbb{E}}_n$  in the proof for  $w_{\text{KCMC}}^*$  above and writing the maximizers of dual problem

$$\sup_{\substack{\eta_{\text{KCMC}} \in \mathbb{R}^D, \\ \eta_f > 0}} -\eta_f \gamma + \eta_{\text{KCMC}}^T \hat{\mathbb{E}}_n[\boldsymbol{\psi}] - \eta_f \hat{\mathbb{E}}_n \left[ f_{T,X}^* \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r}{\eta_f} \right) \right]. \quad (27)$$

as  $\hat{\eta}_f$  and  $\hat{\eta}_{\text{KCMC}}$ , we get

$$\hat{w}_{\text{KCMC}}(y, t, x) \in \left( \frac{1}{p_{\text{obs}}(t|x)} \right) \partial f_{t,x}^* \left( \frac{\hat{\eta}_{\text{KCMC}}^T \boldsymbol{\psi}(t, x) - r(y, t, x)}{\hat{\eta}_f} \right),$$

which proves (23).

For general choice of  $f_{t,x}$ , it is difficult to derive analytical expressions for solutions  $\eta_f^*$  and  $\eta_{\text{KCMC}}^*$ , as well as  $w_{\text{KCMC}}^*$ . However, we can actually obtain their explicit expressions in the case of the box constraints.

**Example 3 (Solutions for box constraints)** *Let us consider the box constraints corresponding to (9) of the conditional f-constraint. Here, for notational simplicity, we omit the subscript of  $a_{\tilde{w}}$  and  $b_{\tilde{w}}$ . For example, we will simply write  $f_{t,x}(\tilde{w}) = I_{[a(t,x), b(t,x)]}(\tilde{w})$ . Then, we can show that*

$$\eta_{\text{KCMC}}^*(t, x) = \left( \frac{\pi(t|x)}{p_{\text{obs}}(t|x)} \right) Q(t, x),$$

where  $Q(t, x)$  is defined as the  $\tau(t, x)$ -quantile of the conditional distribution of  $Y$  given  $T = t$  and  $X = x$  for  $\tau(t, x) := \frac{1 - a(t, x)}{b(t, x) - a(t, x)}$ . From this dual solution, the primal solution can also be recovered using (21) as

$$w_{\text{KCMC}}^*(y, t, x) = \begin{cases} b(t, x) & \text{if } y \leq Q(t, x), \\ a(t, x) & \text{otherwise.} \end{cases}$$

See Appendix C.1 for details.

### 3.2 Specification error

Now we study the specification error of estimator  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi)$ , defined as  $|V_{\text{inf}}^{\text{KCMC}}(\pi) - V_{\text{inf}}^{\text{CMC}}(\pi)|$ . It turns out that it is possible to upper bound the specification error if the dual objective satisfies a Lipschitz condition.

**Theorem 2** *Let  $\|f\|$  denote the  $L^2(\mathcal{T} \times \mathcal{X}, p_{\text{obs}})$  norm for any  $f \in L^2(\mathcal{T} \times \mathcal{X}, p_{\text{obs}})$ . Let  $\Pi_{\psi}$  be the projection operator onto the subspace spanned by  $\{\psi_1, \dots, \psi_D\}$  in  $L^2(\mathcal{T} \times \mathcal{X}, p_{\text{obs}})$ , and let  $(\eta_f^*, \eta_{\text{CMC}}^*)$  be the solution of dual problem (24) for  $V_{\text{inf}}^{\text{CMC}}$ . Additionally, define convex functional  $J : L^2(\mathcal{T} \times \mathcal{X}, p_{\text{obs}}) \rightarrow \mathbb{R}$  corresponding to the negative of dual objective (24) for  $V_{\text{inf}}^{\text{CMC}}$  with fixed  $\eta_f^*$  as*

$$J[h] := \eta_f^* \gamma - \mathbb{E}[h] + \eta_f^* \mathbb{E} \left[ f_{t,x}^* \left( \frac{h(T, X) - r(Y, T, X)}{\eta_f^*} \right) \right].$$

*Let us choose  $\{\psi_1, \dots, \psi_D\}$  such that  $\|\Pi_{\psi} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\| \leq \varepsilon_{\text{spec}}$  for some  $\varepsilon_{\text{spec}} > 0$ . Then, if  $J[h]$  is  $L$ -Lipschitz in neighborhood  $\{h : \|h - \eta_{\text{CMC}}^*\| \leq \varepsilon_{\text{spec}}\}$ , or equivalently, if functional subgradient  $\partial J$  satisfies  $\|\delta J\| \leq L$  for any  $\delta J \in \bigcup_{h: \|h - \eta_{\text{CMC}}^*\| \leq \varepsilon_{\text{spec}}} \partial J[h]$ , we have*

$$|V_{\text{inf}}^{\text{CMC}} - V_{\text{inf}}^{\text{KCMC}}| \leq L \|\Pi_{\psi} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\|.$$

*Moreover, even if  $J[h]$  does not satisfy the  $L$ -Lipschitz condition, if*

$$\eta_{\text{CMC}}^* \in \text{span}(\{\psi_1, \dots, \psi_D\}), \quad (28)$$

*we have no specification error so that*

$$V_{\text{inf}}^{\text{CMC}}(\pi) = V_{\text{inf}}^{\text{KCMC}}(\pi).$$

**Proof** See Appendix C.2. ■

In Theorem 2, we provided the upper bound of the specification error in terms of the Lipschitz constant. A natural follow-up question would be whether it is possible to know such a constant. Indeed, it is possible to calculate the Lipschitz constant in some cases.

**Example 4 (Lipschitz constant for box constraints)** *Consider the same settings as Example 3. Then, if the upper bound of the box constraint satisfies  $b \in L^2(\mathcal{T} \times \mathcal{X}, p_{\text{obs}})$ , the Lipschitz condition for Theorem 2 is satisfied. See Appendix C.3 for details.*

**Example 5 (Lipschitz constant for bounded conditional f-constraint)** *Consider an extension of the previous example where we additionally impose a conditional f-constraint with  $f_{t,x}^0$  so that  $f_{t,x}(\tilde{w}) := I_{[a(t,x), b(t,x)]}(\tilde{w}) + f_{t,x}^0(\tilde{w})$ . The uncertainty set under this constraint can be considered as the intersection of the uncertainty sets of the box constraints and the conditional f-constraint. Interestingly, even for this constraint, we obtain the same Lipschitz constant as the previous example. See Appendix C.4 for details.*

Note that the above examples provide a uniform Lipschitz constant for any policy  $\pi$ . This is in contrast to Theorem 2, which only provide pointwise error bounds for fixed policy  $\pi$  as it assumes that  $\eta_{\text{CMC}}^*$  is fixed. Unfortunately, the uniform Lipschitz constant alone cannot provide a uniform bound on the specification error. Bounding the specification error calls for a uniform bound  $\|\Pi_\psi \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\|$ , but the derivation of such a bound is difficult. Thus, in the policy learning, we have no choice but to optimize  $V_{\text{inf}}^{\text{KCMC}}$ , which is only guaranteed to be lower than  $V_{\text{inf}}^{\text{CMC}}$ .

However, in the policy evaluation, it is possible to provably reduce  $\|\Pi_\psi \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\|$  when we pick the orthogonal functions for kernel conditional moment constraints using the kernel principal component analysis (PCA) (Schölkopf et al., 1997).

**Lemma 3 (Convergence of  $\|\Pi_\psi \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\|$  with kernel PCA)** *Let  $k((t, x), (t', x'))$  be a kernel of feature  $(T, X)$ . Let  $\varphi^{\text{KPCA}}(t, x) := (\varphi_1^{\text{KPCA}}(t, x), \dots, \varphi_D^{\text{KPCA}}(t, x))^T$  be the kernel principal components that approximates the original kernel as  $k((t, x), (t', x')) \approx \varphi^{\text{KPCA}}(t, x)^T \varphi^{\text{KPCA}}(t', x')$  obtained by applying the kernel PCA to the observations of  $\{(T_i, X_i)\}_{i=1}^n$ . Now, suppose the orthogonal functions are chosen as  $\psi := \varphi^{\text{KPCA}}(t, x)$ . Then, if kernel  $k((t, x), (t', x'))$  satisfies a set of regularity conditions<sup>14</sup>, we have*

$$\lim_{D \rightarrow \infty} \text{plim}_{n \rightarrow \infty} \|\Pi_\psi \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\| = 0,$$

where  $\text{plim}$  denotes convergence in probability.

**Proof** See Appendix C.5. ■

Therefore, the kernel PCA can be a useful tool for obtaining an orthogonal basis function set with (close to) zero specification error. When using finite basis functions, manually crafting the basis functions satisfying such a condition is practically difficult. However, the use of kernel PCA allows us to gradually reduce the specification error as we increase the number of basis functions and the sample size, which enables us to achieve consistency under an appropriate choice of kernel thanks to Theorem 2 and 4. In practice, the GIC from Example 8 (or cross-validation) can be used to optimally balance the number of basis functions (i.e., kernel principal components) and the sample size.

Additionally, with the latter statement in Theorem 2 on the no specification bias, we can derive the previously proposed sharp estimator by Dorn and Guo (2022) as a special case.

**Example 6 (Derivation of quantile balancing estimator Dorn and Guo (2022))**  
*Let us consider the box constraints as in Example 3 and choose  $a(t, x)$  and  $b(t, x)$  to be the marginal sensitivity model (8) so that  $\tau(t, x) = \tau = \frac{1}{1+\Gamma}$ . Let  $\hat{Q}(t, x)$  be a solution of  $\tau$ -quantile regression of  $Y$  on  $(T, X)$ . Then, by using a single basis function in the KCMC so that  $D = 1$  and  $\psi_1(t, y) = \left(\frac{\pi(t|x)}{p_{\text{obs}}(t|x)}\right) \hat{Q}(t, x)$ , we recover the original quantile balancing (QB) estimator by (Dorn and Guo, 2022).<sup>15</sup> Especially, if the estimated quantile equals*

14. The formal version of this statement can be found in Appendix C.5.

15. Technically, their estimator employs the linear fractional programming technique from Zhao et al. (2019), but we consider the QB estimator without such a technique here.

the population quantile so that  $\hat{Q}(t, x) = Q(t, x)$ , both KCMC and QB estimators have no specification bias.

Furthermore, suppose quantile estimate  $\hat{Q}(t, x)$  is estimated by a linear quantile regression with  $D$ -dimensional features. Then, we can construct  $D$ -dimensional basis functions for the KCMC using the same features and guarantee that the KCMC estimate of the lower bound is tighter than or equal to that of the QB estimator.

**Proof** See Appendix C.6. ■

**Remark 2 (Comparison of the KCMC estimator and the QB estimator)** *As our estimator generalizes the previous work, our estimator overcomes some drawbacks of the quantile balancing estimators. As discussed in Section 1, the quantile balancing estimator cannot handle policy learning and the  $f$ -constraint. Policy learning is also difficult with the quantile balancing estimator because taking the derivative with respect to policy requires differentiability of the solution of the above linear programming with respect to parameter  $\hat{Q}$ . Moreover, the quantile balancing method is designed only for the box constraints and does not have a proper extension to the  $f$ -sensitivity model (10). In contrast, our estimator of sharper bound  $V_{\text{inf}}^{\text{KCMC}}$  based on the kernel method can naturally handle the above-mentioned generalized cases of sensitivity analysis.*

### 3.3 Consistency of policy evaluation

Now, we study empirical estimator  $\hat{V}_{\text{inf}}^{\text{KCMC}}$  and provide convergence guarantees for policy evaluation and learning. First, we prove the consistency of our estimator for fixed policy  $\pi$  by reducing our problem to the M-estimation (Van de Geer, 2000) using the dual formulation.

#### CONSISTENCY OF POLICY EVALUATION

To align our analysis with the classic framework of M-estimation, let us introduce loss function  $\ell : \Theta \times \mathcal{Z} \rightarrow \mathbb{R}$ , where  $\Theta \subseteq \mathbb{R}^K$  for some  $K$  and  $\mathcal{Z} := \mathcal{Y} \times \mathcal{T} \times \mathcal{X}$ . Then, the following consistency result holds:

**Theorem 4 (Consistency of policy evaluation (informal version))** *Define the parameter space of  $(\eta_f, \eta_{\text{KCMC}})$  as  $\Theta \subseteq \mathbb{R}_+ \times \mathbb{R}^D$ . Further, define  $\theta^* := (\eta_f^*, \eta_{\text{KCMC}}^*)$  as the solution of dual problem (26) for  $V_{\text{inf}}^{\text{KCMC}}(\pi)$  and  $\hat{\theta}_n := (\hat{\eta}_f, \hat{\eta}_{\text{KCMC}})$  as the solution to dual problem (27) for  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi)$ . Define  $\ell : \Theta \times \mathcal{Z} \rightarrow \mathbb{R}$  as*

$$\ell_{\theta}(t, y, x) := \eta_f \gamma - \eta_{\text{KCMC}}^T \psi(t, x) + \eta_f f_{t,x}^* \left( \frac{\eta_{\text{KCMC}}^T \psi(t, x) - r(y, t, x)}{\eta_f} \right)$$

so that it is the negative version of dual objectives (26) and (27) before taking the expectations. Then, if  $\ell_{\theta}(t, y, x)$  satisfies a set of regularity conditions<sup>16</sup>, we have  $\hat{\theta}_n \xrightarrow{P} \theta^*$  and  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi) \xrightarrow{P} V_{\text{inf}}^{\text{KCMC}}(\pi)$ .

**Proof** See Appendix C.7 ■

16. The formal version of this statement can be found in Appendix C.7.

### 3.4 Consistency of policy learning

Here, we provide consistency guarantees for policy learning with the KCMC estimator for a finite dimensional concave policy class and a Vapnik-Chevonenkis (VC) policy class. Again, we take advantage of the reduction to the M-estimation. This simplifies the proof compared to the one in Kallus and Zhou (2021) using the original nested max-min formulation because the max-max formulation we use is simple and unnested maximization, for which the well-studied theory of M-estimation can be immediately applied.

For both proofs, we define a new parameter space and a loss function. Let us define parameter  $\theta = (\beta, \eta_f, \eta_{\text{KCMC}})$ ,  $\eta = (\eta_f, \eta_{\text{KCMC}})$  and its space  $\Theta = \mathcal{B} \times H$  and  $H = \mathbb{R}_+ \times \mathbb{R}^D$ . Define loss function  $\ell : \Theta \times \mathcal{Z} \rightarrow \mathbb{R}$  as

$$\ell_\theta(t, y, x) := \eta_f \gamma - \eta_{\text{KCMC}}^T \psi(t, x) + \eta_f f_{t,x}^* \left( \frac{\eta_{\text{KCMC}}^T \psi(t, x) - \left( \frac{\pi_\beta(t|x)}{p_{\text{obs}}(t|x)} \right) y}{\eta_f} \right) \quad (29)$$

so that it is the negative version of dual objectives (26) and (27) before taking the expectations. Define also  $\theta^* := (\eta_f^*, \eta_{\text{KCMC}}^*, \beta^*)$  so that  $\beta^* \in \arg \max_{\beta \in \mathcal{B}} V_{\text{inf}}^{\text{KCMC}}(\pi_\beta)$  and  $\eta^* := (\eta_f^*, \eta_{\text{KCMC}}^*)$  is the solution of dual problem (26) for  $V_{\text{inf}}^{\text{KCMC}}$  at policy  $\pi_{\beta^*}$ . Similarly, define  $\hat{\theta}_n := (\hat{\eta}_f, \hat{\eta}_{\text{KCMC}}, \hat{\beta})$  so that  $\hat{\beta} \in \arg \max_{\beta \in \mathcal{B}} \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_\beta)$  and  $\hat{\eta} := (\hat{\eta}_f, \hat{\eta}_{\text{KCMC}})$  is the solution of dual problem (27) for  $\hat{V}_{\text{inf}}^{\text{KCMC}}$  at policy  $\pi_{\hat{\beta}}$ .

#### CONCAVE POLICY LEARNING

With the above definitions, we can now show the consistency of policy learning with a concave policy class.

**Theorem 5 (Consistency of concave policy learning (informal version))** *Assume concave policy class  $\{\pi_\beta(t|x) : \beta \in \mathcal{B}\}$  with convex parameter space  $\mathcal{B}$  satisfying that  $\beta \mapsto \pi_\beta(t|x)y$  is concave for any  $y \in \mathcal{Y}$ ,  $t \in \mathcal{T}$  and  $x \in \mathcal{X}$ . Then, if loss function (29) satisfies a set of regularity conditions<sup>17</sup>, we have  $\hat{\theta}_n \xrightarrow{p} \theta^*$  and  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) \xrightarrow{p} V_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*})$ .*

**Proof** See Appendix C.8. ■

An example of concave policy is mixed policy  $\pi_\beta(t|x) := \sum_k \beta_k \pi_k(t|x)$  for  $\sum_k \beta_k = 1$ ,  $\beta_k \geq 0$ . Indeed, policy learning with such a concave policy class is concave; therefore, the globally optimal policy can be found by convex optimization algorithms.

#### VAPNIK-CHEVONENKIS POLICY LEARNING

Now, let us discuss the consistency of policy learning with a VC policy class. While the proof for VC policy class is more involved than the concave policy class, we can prove a similar consistency guarantee for policy learning with a VC policy class:

**Theorem 6 (Consistency of VC policy learning (informal version))** *Let  $\{\pi_\beta(t|x) : \beta \in \mathcal{B}\}$  be a class of policies that is VC. Under a set of regularity conditions<sup>18</sup>, we have  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) \xrightarrow{p} V_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*})$ .*

17. The formal version of this statement can be found in Appendix C.8.

18. The formal version of this statement can be found in Appendix C.9.

**Proof** See Appendix C.9. ■

Here, we adopt the same definition of the VC function class as in Van de Geer (2020, Definition 6.4.1), where they define it as the class of functions whose subgraphs collectively form a VC set.

### 3.5 Asymptotic normality of policy evaluation

Now, we consider the asymptotic normality in the policy evaluation settings.

**Theorem 7 (Asymptotic normality (informal version))** *Under a set of regularity conditions<sup>19</sup>, we have*

$$\sqrt{n} \left[ \hat{\mathbb{E}}_n \ell_{\hat{\theta}}(Z) - \hat{\mathbb{E}}_n \ell_{\theta^*}(Z) \right] \xrightarrow{p} 0,$$

and

$$\sqrt{n} \left[ V(\hat{\theta} - \theta^*) + \hat{\mathbb{E}}_n \delta \ell_{\theta^*}(Z) \right] \xrightarrow{p} 0.$$

Therefore,

$$\sqrt{n}[\hat{\theta} - \theta^*] \xrightarrow{p} V^{-1} \sqrt{n} \hat{\mathbb{E}}_n [\delta \ell_{\theta^*}(Z)] \xrightarrow{d} N(0, V^{-1} J V^{-1}),$$

where  $V, J$  are defined as  $V := \nabla_{\theta} \nabla_{\theta}^T \mathbb{E} \ell_{\theta^*}(Z)$  and  $J := \mathbb{E} [\delta \ell_{\theta^*}(Z) \delta \ell_{\theta^*}(Z)^T]$ . Convergence symbol  $\xrightarrow{d}$  implies the convergence in distribution.

**Proof** See Appendix C.10. ■

Now, let us consider a few applications of the above asymptotic result.

**Example 7 (Confidence interval)** *Under Condition 1 and Condition 3, we know that empirical objective  $\sqrt{n} \hat{\mathbb{E}}_n \ell_{\hat{\theta}}$  has the same asymptotic distribution as  $\sqrt{n} \mathbb{E}_n \ell_{\theta^*}$ , which is  $N(\mathbb{E}[\ell_{\theta^*}(Z)], \mathbb{V}[\ell_{\theta^*}(Z)])$ . Therefore, the confidence interval of the lower bound  $V_{\text{inf}}^{\text{KCMC}}(\pi) = -\mathbb{E}[\ell_{\theta^*}(Z)]$  with significance level  $\alpha$  is*

$$C_n^{\text{V}_{\text{inf}}^{\text{KCMC}}} = [C_n^-, C_n^+]$$

for

$$C_n^{\pm} = \mathbb{E}[\ell_{\theta^*}(Z)] \pm \sqrt{\mathbb{V}[\ell_{\theta^*}(Z)]/n} \cdot \Phi^{-1}(1 - \alpha/2),$$

where  $\Phi(s)$  defined as the inverse of the cumulative density function of standard normal distribution  $\Phi(s) := \int_{-\infty}^s \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}t^2) dt$ . In practice, we can estimate  $\mathbb{E}[\ell_{\theta^*}(Z)]$  and  $\mathbb{V}[\ell_{\theta^*}(Z)]$  with the sample averages using the M-estimator in place of the true parameter as  $\hat{\mathbb{E}}_n[\ell_{\hat{\theta}}(Z)]$  and  $\hat{\mathbb{V}}_n[\ell_{\hat{\theta}}(Z)]$ .

**Example 8 (Generalized information criterion)** *In the previous example, we approximated true lower bound  $\mathbb{E}[\ell_{\theta^*}(Z)]$  with  $\hat{\mathbb{E}}_n[\ell_{\hat{\theta}}(Z)]$ . Though this approximation is a correct first*

<sup>19</sup> The formal version of this statement can be found in Appendix C.10.

order approximation, we can consider its second order correction as follows, as discussed in Appendix C.11:

$$n\hat{\mathbb{E}}_n[\ell_{\hat{\theta}}(Z) - \ell_{\theta^*}(Z)] \stackrel{p}{\rightarrow} -\frac{1}{2}\sqrt{n}(\hat{\theta} - \theta^*)^T V\sqrt{n}(\hat{\theta} - \theta^*).$$

Thus, the bias of  $\hat{\mathbb{E}}[\ell_{\hat{\theta}}(Z)]$  can be written as  $-\mathbb{E}\left[\frac{1}{2}(\hat{\theta} - \theta^*)^T V(\hat{\theta} - \theta^*)\right] = -\frac{1}{2n}\text{tr}[V^{-1}J]$  since  $n \cdot \mathbb{E}\left[(\hat{\theta} - \theta^*)(\hat{\theta} - \theta^*)^T\right] \stackrel{p}{\rightarrow} V^{-1}JV^{-1}$ . Using this second order bias correction, we can obtain the generalized information criterion (GIC) (Konishi and Kitagawa, 2008)<sup>20</sup> of the lower bound as

$$\hat{V}_{\text{inf,GIC}}^{\text{KCMC}}(\pi) = \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi) - \frac{1}{2n}\text{tr}[\hat{V}^{-1}\hat{J}],$$

where  $\hat{J} := \hat{\mathbb{E}}_n[\delta\ell_{\hat{\theta}}(Z)\delta\ell_{\hat{\theta}}(Z)^T]$  and  $\hat{V} := \hat{\mathbb{E}}_n[\nabla_{\theta}\nabla_{\theta}^T\ell_{\hat{\theta}}(Z)]$  when  $\theta \mapsto \ell_{\theta}(z)$  is twice differentiable at  $\theta = \theta^*$ . When the twice differentiability does not hold for the loss function, we need to derive the analytical form of Hessian  $V$  and construct its estimator. Indeed, the loss function is not pointwise differentiable in the case of box constraints, and the analytical form of its Hessian needs to be derived as in Appendix D.

**Example 9 (Confidence interval with second order bias correction)** Using the second order bias correction above, we can obtain a new confidence interval of  $V_{\text{inf}}^{\text{KCMC}}(\pi) = -\mathbb{E}[\ell_{\theta^*}(Z)]$  as

$$\tilde{C}_n^{\text{KCMC}} = [\tilde{C}_n^-, \tilde{C}_n^+]$$

for

$$\tilde{C}_n^{\pm} = \mathbb{E}[\ell_{\theta^*}(Z)] - \frac{1}{2n}\text{tr}[V^{-1}J] \pm \sqrt{\mathbb{V}[\ell_{\theta^*}(Z)]/n} \cdot \Phi^{-1}(1 - \alpha/2).$$

We examine the benefit of this second order bias correction in our numerical experiments in Section 4.3.

## 4. Numerical Experiments

In this section, we present the results of our numerical experiments. Our experiments demonstrate the applicability of the KCMC estimator in various settings such as sensitivity analysis with a continuous (treatment) action space, the new f-sensitivity model, policy learning, construction of confidence interval, and model selection.

### 4.1 Experimental settings

In all the experiments except Section 4.2, we use synthetic data with a binary action space sampled from the following data generating process:

$$X \sim \mathcal{N}(\mu_x, \mathbf{I}_5),$$

20. Note that the original GIC corrects bias term  $\hat{\mathbb{E}}_n[\ell_{\hat{\theta}}(Z)] - \mathbb{E}[\ell_{\hat{\theta}}(Z)] \stackrel{p}{\rightarrow} -\frac{1}{n}\text{tr}[V^{-1}J]$ , which has multiplier  $\frac{1}{n}$  instead of  $\frac{1}{2n}$  of our bias adjustment term of the lower bound. This is because the GIC corrects both underestimation bias of training risk  $\mathbb{E}(\hat{\mathbb{E}}_n[\ell_{\hat{\theta}}(Z)]) - \mathbb{E}[\ell_{\theta^*}(Z)]$  and generalization error  $\mathbb{E}[\ell_{\theta^*}(Z)] - \mathbb{E}[\ell_{\hat{\theta}}(Z)]$  while our estimator only corrects the former.

$$\begin{aligned}
 Y(0)|X &\sim \mathcal{N}(\beta_{x,0}^T X + \beta_{\text{const},0}, 1), \\
 Y(1)|X &\sim \mathcal{N}(\beta_{x,1}^T X + \beta_{\text{const},1}, 1), \\
 T|X &\sim \text{Bern}\left(\frac{1}{1 + \exp(-\beta_t^T X)}\right), \\
 Y &= Y(T),
 \end{aligned} \tag{30}$$

where

$$\begin{aligned}
 \mu_x &= (-1, 0.5, -1, 0, -1)^T, \\
 \beta_{x,0} &= (0, 0.5, -0.5, 0, 0)^T, \\
 \beta_{x,1} &= (-1.5, 1.5, -2, 1, 0.5)^T, \\
 \beta_{\text{const},0} &= 2.5, \\
 \beta_{\text{const},1} &= 0.5, \\
 \beta_t &= (0, 0.75, -0.5, 0, -1)^T.
 \end{aligned}$$

Though this model is not confounded at all, it has a partially analytical solution of the policy value lower bound for box constraints. This semi-analytical lower bound can be computed using the Monte Carlo method, in a similar manner to the synthetic data in Dorn and Guo (2022, Corollary 3.). In the policy evaluation, we use logistic policy  $\pi(T = 1|X) = 1 / (1 + \exp(-\beta_\pi^T X))$ , where  $\beta_\pi = (1, 0.5, -0.5, 0, 0)^T$ , and we consider box constraints  $\Gamma^{-1} \leq \frac{p_{\text{obs}}(t|x)(1-\pi_{\text{base}}(t|x,u))}{(1-p_{\text{obs}}(t|x))\pi_{\text{base}}(t|x,u)} \leq \Gamma$  unless otherwise specified.

Similarly to this, we created a variation of the above synthetic data with continuous action space. In the continuous version, we replaced the fourth line and the fifth line of data generating process for the binary data (30) with

$$\begin{aligned}
 T|X &\sim \mathcal{N}\left(\frac{1}{1 + \exp(-\beta_t^T X)}, 1\right), \\
 Y(t) &:= (1 - t)Y(0) + tY(1), \\
 Y &= Y(T)
 \end{aligned}$$

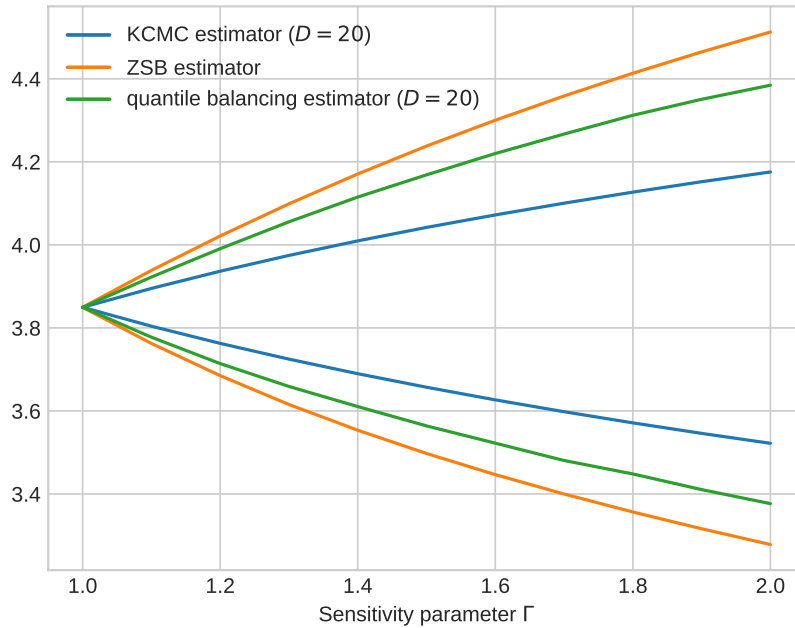
so that reward  $Y$  linearly depends on treatment  $T$  given  $Y(0)$  and  $Y(1)$ . We can think of this model as the case where partial treatment (i.e.,  $0 < T < 1$ ) is possible, and its outcome becomes an interpolation of no treatment  $T = 0$  and full treatment  $T = 1$ . In the policy evaluation, we use Gaussian policy  $\pi(t|X) = \mathcal{N}(t; \beta_\pi^T X, 0.25)$ . For the continuous synthetic data, we consider box constraints  $\Gamma^{-1} \leq \frac{\pi_{\text{base}}(t|x,u)}{p_{\text{obs}}(t|x)} \leq \Gamma$ .

We also included the same real-world data as Dorn and Guo (2022), which is 668 subsamples of data from the 1966-1981 National Longitudinal Survey (NLS) of Older and Young Men. These subsamples consist of the 1978 cross-section of Young Men who are craftsmen or laborers and are not enrolled in school. For this data, we consider box constraints  $\Gamma^{-1} \leq \frac{p_{\text{obs}}(t|x)(1-\pi_{\text{base}}(t|x,u))}{(1-p_{\text{obs}}(t|x))\pi_{\text{base}}(t|x,u)} \leq \Gamma$ .

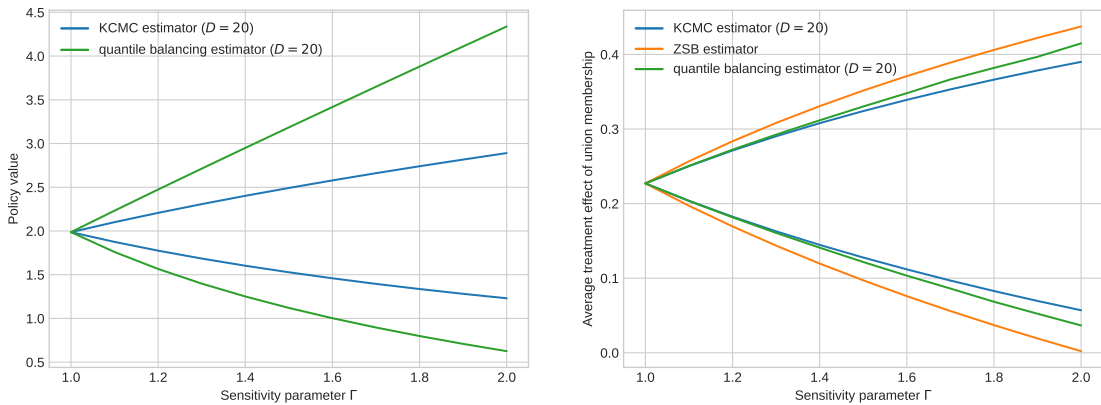
Conditional probability  $p_{\text{obs}}(t|x)$  needed to construct the estimators was calculated from the true data generating process in the case of synthetic data, and it was estimated from the data using the logistic regression with covariate  $X$  in the case of the real-world example.

To solve the convex programming involved in the above estimators, we used MOSEK (ApS, 2019) and ECOS (Domahidi et al., 2013) through the API of CVXPY (Diamond and Boyd, 2016). Lastly, the experimental code necessary to reproduce the following results will be made fully available at <https://github.com/kstoneriv3/confounding-robust-inference>.

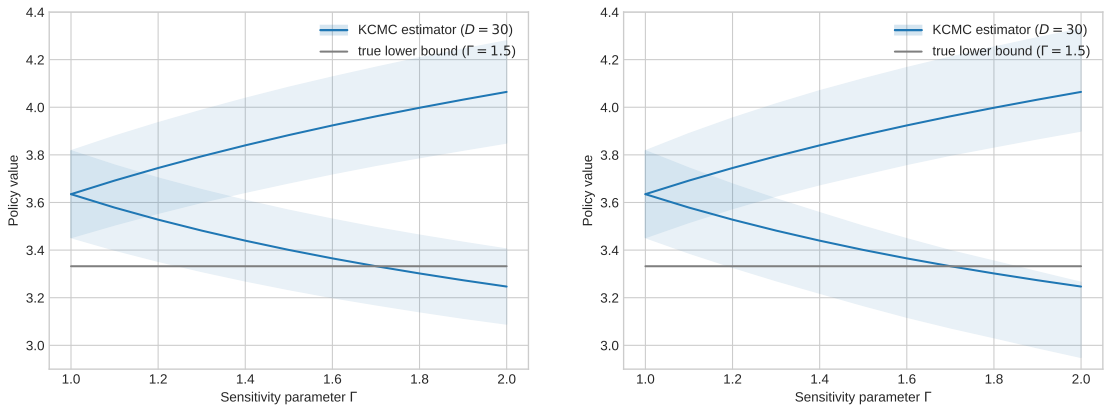
## 4.2 Comparison of KCMC estimator to baselines



**Figure 2:** Estimated upper and lower bounds of policy value using different sensitivity parameter  $\Gamma$  for the synthetic data of sample size 1000 with binary action space.



**Figure 3:** Estimated upper and lower bounds of policy value using different sensitivity parameters. Synthetic data of sample size 1000 with continuous action space (left) and the NLS data of sample size 668 with a binary action space (right) are used.



**Figure 4:** Estimated upper and lower bounds of policy value with 95% confidence interval with (right) and without (left) second-order correction. Solid lines and the bands surrounding them indicate the point estimate (both without second order correction) and its confidence interval.

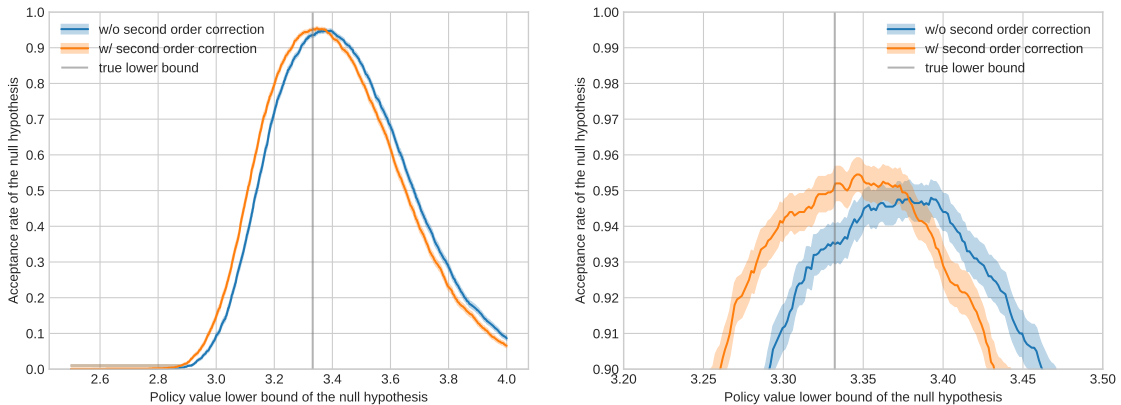
Figure 2 shows the upper and lower bounds of the policy value estimate for binary synthetic data for KCMC estimator (18), ZSB estimator (Zhao et al., 2019), and the quantile balancing (QB) estimator (Dorn and Guo, 2022)<sup>21</sup> under different confounding levels  $\Gamma$ . Here, fractional programming is not used in the ZSB estimator and the quantile balancing estimator in order to match the policy value with the KCMC estimator under no confounding (i.e.,  $\Gamma = 1.0$ ). The feature vectors for the QB estimator and the KCMC estimator were chosen so that they represent the Example 6, where the first stage of the QB estimator solves the modified version of the dual problem for the KCMC estimator. As discussed in Example 6, the QB estimator is always no tighter than the KCMC estimator under such a choice of the feature vector, and it can clearly be observed in this example. Indeed, we can see that the tightness of the bounds loosens as sensitivity parameter  $\Gamma$  gets larger. When compared to the sharp estimators (KCMC and QB), the ZSB estimator has looser bounds.

We also present the bounds of the policy values for the synthetic data with continuous action space and real-world data in Figure 3. Please note that in the case of synthetic data with continuous action space, the ZSB estimator cannot be calculated as it requires that the action space be discrete and finite. In both cases, the tightness of the bounds depends on the sensitivity parameter and the type of estimators in the same way as the case of the binary synthetic data.

### 4.3 Confidence interval

Figure 4 shows the confidence intervals of upper and lower bounds of policy values discussed in Example 7 and 9. Since the second order debiasing term corrects the overly optimistic bounds due to the overfitting, the confidence intervals with the second order correction provide more pessimistic bounds.

21. For ZSB estimator, use the empirical version of (16). For both ZSB and the QB estimator, we do not use the linear fractional programming technique (Zhao et al., 2019) to match the estimator values under no confounding for the purpose of comparison.



**Figure 5:** Acceptance rate of the null hypothesis under different null hypotheses with Tan’s box constraints ( $\Gamma = 1.5$ ) with significance level  $\alpha = 0.05$ . The plot on the left is in the original size and the one on the right is its zoomed version.

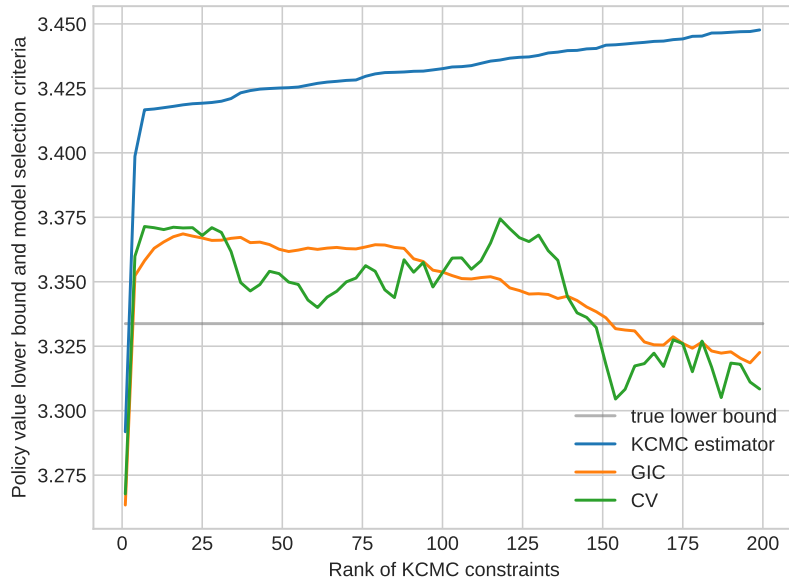
In Figure 5, we show the acceptance rate of the null hypotheses with different values of true lower bound. The acceptance rate is calculated as the coverage rate of the confidence intervals for the synthetic data of sample size 2000, simulated 2000 times. As noted earlier, this synthetic data has an analytically tractable true lower bound and it is indicated by the grey vertical line. As can be expected, acceptance rates of both confidence intervals (with and without second order bias correction) reach their peaks near the true lower bound. However, the confidence interval without second order bias correction has its peaks shifted towards the optimistic direction due to the overfitting in the dual problem and exhibits some level of overrejection at the true null hypothesis compared to its counterpart with the bias correction.

#### 4.4 Model Selection

Here, we show an example use of the GIC discussed in Example 8 for rank selection of the KCMC. As the KCMC is a finite dimensional approximation of the infinite dimensional constraints, we have the motivation to increase the rank of the KCMC estimator. On the other hand, the original KCMC lower bound increases monotonously as we increase the rank of KCMC constraints. We can even recover the inverse probability weighting estimator<sup>22</sup> by making the KCMC full-rank, but this nullifies the confounding robustness, so we would like to avoid using the excessively high rank. This necessitates systematic procedures for selecting the appropriate rank of the KCMC. Fortunately, our dual problem can be interpreted as a standard empirical risk minimization, where the rank selection can be interpreted as the model selection where the cross-validation and the GIC can be used as the selection criteria.

In Figure 6, we plotted the KCMC lower bound estimator, its GIC, and its cross validation. While the original estimator increases monotonically as the rank of the KCMC increases, the GIC and the cross validation start to decline when the rank of the constraints becomes

22. More precisely, the inverse probability weighting estimator that uses the  $p_{\text{obs}}(t|x)$  in place of the true propensity score.



**Figure 6:** Estimated policy value lower bound for box constraints ( $\Gamma = 1.5$ ), its generalized information criterion, and its 10-fold cross validation (CV) for the synthetic data ( $n = 1000$ ) with binary action space.

excessively high. In this example, both the GIC and the cross validation seem to prefer ranks around 20, though the cross validation tends to be more noisy and also prefers ranks of around 120.

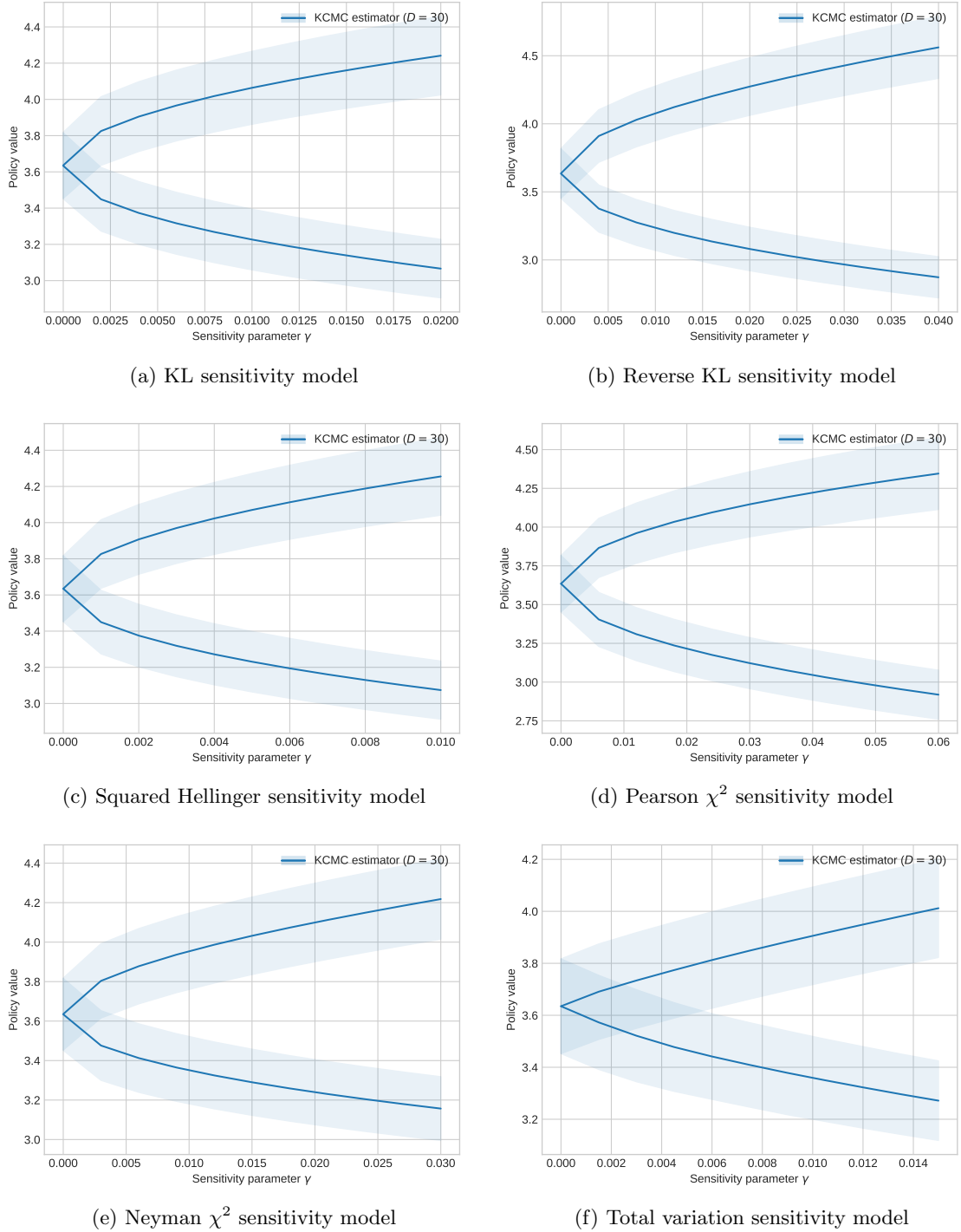
#### 4.5 f-sensitivity model

In Figure 7, we present the upper/lower bounds and confidence intervals of the policy value given by the KCMC estimators of f-sensitivity models. Similarly to the example of box constraints, we can see that the f-sensitivity models can produce a continuous control of the level of confounding by the sensitivity parameter. We can also confirm that the confidence intervals behave quite similarly to the case of box constraints.

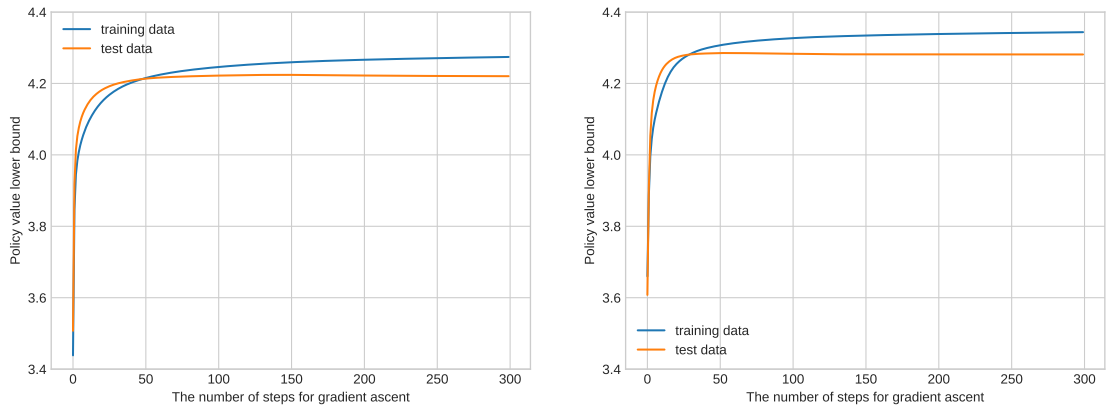
#### 4.6 Policy learning

Finally, we consider the max-min policy learning with the KCMC estimator. In the policy learning, we use the gradient ascent on the lower bound estimator similarly to Kallus and Zhou (2018, 2021).<sup>23</sup> As a baseline, we also implemented policy learning with the ZSB estimator using the same dataset and initial policy. In Figure 8, we present the learning curves of both estimators. As can be expected, after the max-min policy learning, the KCMC estimator achieves a higher lower bound than the ZSB estimator, both on the training data and the test data.

23. By Danskin’s theorem (Danskin, 1966), we can calculate the gradient for outer maximization as the gradient at the solution of the inner maximization problem.



**Figure 7:** Estimated upper and lower bounds of policy value and their confidence intervals using  $f$ -sensitivity models. The synthetic data of sample size 4000 with binary action space is used.



**Figure 8:** Learning curves of the max-min policy learning of logistic policy by the ZSB estimator (left) and the KCMC estimator (right). Both training and test data are binary synthetic data of sample size 1000. The training curve and the test curve represent the values of each estimator for the data.

## 5. Conclusion

In this paper, we proposed a novel framework of the sensitivity analysis using kernel conditional moment constraints. Together with the newly proposed conditional  $f$ -constraints, this approach provides a unified formulation of sensitivity analysis that encompasses a broad class of existing models and extends its applicability to previously intractable settings. From a theoretical perspective, we exploited convexity properties to recast our estimation procedure as an empirical risk minimization problem and established consistency guarantees for both policy evaluation and policy learning using standard tools from M-estimation theory. Finally, we demonstrated the practical effectiveness of the proposed kernel conditional moment constraints through numerical experiments, including several novel extensions of sensitivity analysis.

# Appendix

## Appendix A. Notations

| notation                              | meaning   |
|---------------------------------------|---|
| $Y \in \mathcal{Y}$                   | reward (outcome)  |
| $T \in \mathcal{T}$                   | action (treatment)  |
| $X \in \mathcal{X}$                   | context   |
| $U \in \mathcal{U}$                   | unobserved confounder   |
| $\pi_{\text{base}}(t x, u)$           | unknown (and confounded) base policy from which the data is generated   |
| $\pi(t x)$                            | target policy for policy evaluation (or learning)   |
| $p_{\text{obs}}(y, t, x)$             | marginal distribution of $(Y, T, X)$ under data generating process (1)  |
| $\mathbb{E}, \mathbb{E}_{\text{obs}}$ | expectation under data generating process (1)   |
| $\hat{\mathbb{E}}_n$                  | empirical average over $n$ samples generated from (1)   |
| $\mathcal{E}$                         | uncertainty set of $\pi_{\text{base}}(t x, u)$  |
| $f_{t,x}(v)$                          | a convex function w.r.t. $v$ for fixed $t, x$ satisfying $f_{t,x}(1) = 1$   |
| $f_{t,x}^*(u)$                        | convex conjugate of $f_{t,x}(v)$  |
| $w(y, t, x)$                          | reparametrization $w(y, t, x) = \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot X, U)} \left[ \frac{1}{\pi_{\text{base}}(T X, U)}   Y = y, T = t, X = x \right]$ |
| $\tilde{w}(y, t, x)$                  | reparametrization $\tilde{w}(y, t, x) = p_{\text{obs}}(t x)w(y, t, x)$  |
| $\mathcal{W}$                         | uncertainty set of $w(y, t, x)$ , such as $\mathcal{W}_{f_{t,x}}^{\text{KCMC}}$   |
| $\hat{\mathcal{W}}$                   | uncertainty set of $w(y, t, x)$ defined using empirical samples, such as $\hat{\mathcal{W}}_{f_{t,x}}^{\text{KCMC}}$  |
| $V(\pi)$                              | value of policy $\pi(t x)$ under confounded data generating process (2)   |
| $V_{\text{inf}}(\pi)$                 | lower bound of the value of policy $\pi(t x)$ under confounded data generating process (2)  |
| $\hat{V}_{\text{inf}}(\pi)$           | estimator of $V_{\text{inf}}(\pi)$ using empirical samples  |
| $\psi_d(t, x)$                        | $d$ -th orthogonal function of kernel conditional moment constraints  |
| $\varphi_d^{\text{KPCA}}(t, x)$       | $d$ -th principal component of kernel PCA fitted to empirical data  |
| $\varphi_d^*(t, x)$                   | $d$ -th principal component of kernel PCA fitted to population (i.e. infinite data)   |
| $\lambda_d^*$                         | the eigenvalue of $\varphi_d^*(t, x)$   |
| $n$                                   | the number of samples   |
| $D$                                   | the number of kernel conditional moment constraints   |
| $e(y, t, x)$                          | reparametrization $e(y, t, x) := p_{\text{obs}}(t x)w(y, t, x) - 1$   |
| $r(y, t, x)$                          | reparametrization $r(y, t, x) = \left( \frac{\pi(t x)}{p_{\text{obs}}(t x)} \right) \cdot y$  |
| $\ell_\theta(z)$                      | loss function of dual problems (26) and (27).   |
| $\partial_\theta \ell_\theta(z)$      | subgradient of the loss function w.r.t. $\theta$  |
| $Z$                                   | shorthand notation for observable variables $(Y, T, X)$   |
| $\theta \in \Theta$                   | parameters of dual loss $(\eta_f, \eta_{\text{KCMC}})$ or $(\eta_f, \eta_{\text{KCMC}}, \beta)$   |
| $\theta^*$                            | minimizer of population risk $\mathbb{E} \ell_\theta(Z)$  |
| $\hat{\theta}_n$                      | minimizer of empirical risk $\hat{\mathbb{E}}_n \ell_\theta(Z)$   |
| $\Theta_\varepsilon$                  | $\varepsilon$ -neighborhood of $(\eta^*(\beta), \beta)$ of the partially optimal parameters in policy learning  |

**Table 1:** List of common notations used throughout the paper.

## Appendix B. An Alternative f-Sensitivity Model by Jin et al. (2022)

Jin et al. (2022) proposed a similar uncertainty set that approximates the condition

$$D_f [p(Y(1)|X, T = 1) || p(Y(1)|X, T = 0)] \leq \gamma \text{ almost surely,}$$

where variable  $Y(1)$  is the potential outcome variable for treatment  $T = 1$  in Rubin's potential outcome framework (Rubin, 2005) with binary treatment. Under the assumption of unconfoundedness, the potential outcome variable  $Y(1)$  must satisfy  $Y(1) \perp\!\!\!\perp T|X$ , and thus it must satisfy  $D_f [p(Y(1)|X, T = 1) || p(Y(1)|X, T = 0)] = 0$  almost surely with respect to  $p_{\text{obs}}$ . Their sensitivity model can be interpreted as a relaxation of this assumption by allowing the violation of it up to  $\gamma$ . To highlight the difference in modeling paradigm, our f-sensitivity model follows the same modeling framework as Tan (2006), which takes into account the difference between observational policy  $p_{\text{obs}}(t|x)$  and underlying confounded policy  $\pi_{\text{base}}(t|x, u)$ . On the other hand, the model by Jin et al. (2022) considers the distributional shift between observation  $Y(1)|X = x, T = 1$  and counterfactual  $Y(1)|X = x, T = 0$ , and therefore, their modeling approach is different from Tan (2006) and its extension.

## Appendix C. Omitted Proofs

### C.1 Details of Example 3

Here, we consider the box constraints corresponding to (9) of the conditional f-constraint. For function  $f_{t,x}(\tilde{w}) = I_{[a(t,x), b(t,x)]}(\tilde{w})$ , we can derive its conjugate and its subgradient as

$$f_{t,x}^*(v) = \begin{cases} a(t,x)v & \text{if } v < 0, \\ 0 & \text{if } v = 0, \\ b(t,x)v & \text{if } v > 0, \end{cases}$$

and

$$\partial f_{t,x}^*(v) = \begin{cases} a(t,x) & \text{if } v < 0, \\ [a(t,x), b(t,x)] & \text{if } v = 0, \\ b(t,x) & \text{if } v > 0. \end{cases}$$

Substituting the above expression of  $\partial f_{t,x}^*$  in the first order condition of (25), we can derive more explicit expression

$$\begin{aligned} 1 &= \mathbb{E} [\partial f_{T,X}^*(\eta_{\text{CMC}} - r) | T = t, X = x] \\ &= \mathbb{P}(r < \eta_{\text{CMC}}(t, x)) \cdot b(t, x) + \mathbb{P}(r = \eta_{\text{CMC}}(t, x)) \cdot [a(t, x), b(t, x)] + \mathbb{P}(r > \eta_{\text{CMC}}(t, x)) \cdot a(t, x). \end{aligned}$$

Here, we used the box constraints' property  $a(t, x) \leq 1 \leq b(t, x)$ , which follows from the requirement that  $f_{t,x}$  must satisfy  $f_{t,x}(1) = 0$  for any  $t \in \mathcal{T}$  and  $x \in \mathcal{X}$ . Assuming that the conditional distribution of  $r(Y, T, X)$  given  $T = t$  and  $X = x$  yields continuous distribution for any  $t \in \mathcal{T}$  and  $x \in \mathcal{X}$ , the second subgradient condition becomes

$$1 = \mathbb{P}(r \leq \eta_{\text{CMC}}(t, x)) \cdot b(t, x) + \mathbb{P}(r > \eta_{\text{CMC}}(t, x)) \cdot a(t, x).$$

This implies that  $\mathbb{P}(r \leq \eta_{\text{CMC}}(t, x)) = \frac{1-a(t,x)}{b(t,x)-a(t,x)} =: \tau(t, x)$ , and therefore,

$$\eta_{\text{CMC}}^*(t, x) = \left( \frac{\pi(t|x)}{p_{\text{obs}}(t|x)} \right) Q(t, x),$$

where  $Q(t, x)$  was defined as the  $\tau(t, x)$ -quantile of the conditional distribution of  $Y$  given  $T = t$  and  $X = x$ . From this dual solution, the primal solution can also be recovered using (21) as

$$w_{\text{CMC}}^*(y, t, x) = \begin{cases} b(t, x) & \text{if } y \leq Q(t, x), \\ a(t, x) & \text{otherwise.} \end{cases}$$

## C.2 Proof of Theorem 2

**Proof** As  $\Pi_{\psi} \eta_{\text{CMC}}^*$  is on the subspace spanned by  $\{\psi_1, \dots, \psi_D\}$ , we can take  $\eta_{\text{KCMC}}^* \in \mathbb{R}^D$  such that  $\eta_{\text{KCMC}}^*{}^T \psi = \Pi_{\psi} \eta_{\text{CMC}}^*$ . Then, due to the fundamental theorem of calculus, we have

$$\begin{aligned} J[\eta_{\text{KCMC}}^*{}^T \psi] &= J[\eta_{\text{CMC}}] + \int_{\eta_{\text{CMC}}^*}^{\eta_{\text{KCMC}}^*{}^T \psi} \sup_{\delta J \in \partial J[h]} \langle \delta J, dh \rangle \\ &= J[\eta_{\text{CMC}}] + \int_0^1 \sup_{\delta J \in \partial J[\eta_{\text{CMC}} + s \Delta \eta]} \langle \delta J, \Delta \eta \rangle ds \\ &\leq J[\eta_{\text{CMC}}] + \int_0^1 \sup_{\delta J \in \partial J[\eta_{\text{CMC}} + s \Delta \eta]} \|\delta J\| \cdot \|\Delta \eta\| ds \\ &\leq J[\eta_{\text{CMC}}] + L \|\Delta \eta\| \end{aligned}$$

for  $\Delta \eta := \eta_{\text{KCMC}}^*{}^T \psi - \eta_{\text{CMC}} = \Pi_{\psi} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*$ . As  $V_{\text{inf}}^{\text{CMC}} = -J[\eta_{\text{CMC}}^*]$  and  $V_{\text{inf}}^{\text{KCMC}} \geq \sup_{\eta_{\text{KCMC}}^* \in \mathbb{R}^D} \{-J[\eta_{\text{KCMC}}^*{}^T \psi]\} \geq -J[\eta_{\text{KCMC}}^*{}^T \psi]$ , we get

$$V_{\text{inf}}^{\text{KCMC}} \leq V_{\text{inf}}^{\text{CMC}} \leq V_{\text{inf}}^{\text{KCMC}} + L \|\Pi_{\psi} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\|.$$

To prove the latter statement on no specification error, we take multiplier  $\eta_{\text{KCMC}}^*$  that satisfies  $\eta_{\text{CMC}}^* = \eta_{\text{KCMC}}^*{}^T \psi$ . Then, we can see that dual problem (26) for  $V_{\text{inf}}^{\text{KCMC}}$  can be considered as the restricted version of dual problem (24) for  $V_{\text{inf}}^{\text{CMC}}$  where  $\eta_{\text{CMC}}$  is constrained to the subspace spanned by  $\{\psi_1, \dots, \psi_D\}$ . Therefore, as restricted solution  $(\eta_{\text{KCMC}}^*, \eta_f^*)$  achieves the same value as the solution of the non-restricted problem, it is clearly a solution of restricted problem (26). Finally, owing to the strong duality, we can calculate the values of  $V_{\text{inf}}^{\text{CMC}}$  and  $V_{\text{inf}}^{\text{KCMC}}$  by the values of the dual problems, which implies  $V_{\text{inf}}^{\text{CMC}} = V_{\text{inf}}^{\text{KCMC}}$ . ■

## C.3 Details of Example 4

As discussed previously, in the case of box constraints,  $\partial f^*$  is bounded so that  $\partial f_{t,x}^* \subseteq [a, b]$ . Here, we omitted the subscripts and the arguments as  $a = a_{\bar{w}}(t, x)$  and  $b = b_{\bar{w}}(t, x)$  for notational simplicity. As  $a$  and  $b$  satisfy  $a \leq 1 \leq b$ , due to requirement  $f_{t,x}(1) = 0$ , we get

the following explicit formula for the Lipschitz constant:

$$\begin{aligned}
 L_{a,b} &= \sup \left\{ \|\delta J\| : \delta J \in \bigcup_{h: \|h - \eta_{\text{CMC}}^*\| \leq \varepsilon_{\text{spec}}} \partial J[h] \right\} \\
 &= \sup \left\{ \|\delta J\| : \delta J \in \bigcup_{h: \|h - \eta_{\text{CMC}}^*\| \leq \varepsilon_{\text{spec}}} \mathbb{E} \left[ \partial f_{T,X}^* \left( \frac{h-r}{\eta_f} \right) \middle| T=t, X=x \right] - 1 \right\} \\
 &\leq \sup \left\{ \|\delta J\| : \delta J \in L^2(\mathcal{T} \times \mathcal{X}, p_{\text{obs}}), a(t,x) - 1 \leq \delta J(t,x) \leq b(t,x) - 1 \right\} \\
 &= \left( \mathbb{E} |\max\{1 - a(T, X), b(T, X) - 1\}|^2 \right)^{1/2}.
 \end{aligned}$$

As we know  $0 \leq a \leq b$ , we can see that  $b \in L^2(\mathcal{T} \times \mathcal{X}, p_{\text{obs}})$  is the sufficient condition to have finite Lipschitz constant  $L_{a,b}$ .

#### C.4 Details of Example 5

To show that we can still apply the same argument as Example 4, we can calculate the conjugate and the subgradient of the conjugate as

$$f_{t,x}^*(v) = \sup_{u \in [a,b]} \{vu - f_{t,x}^0(u)\} = \begin{cases} av - f_{t,x}^0(a) & \text{if } v < \partial f_{t,x}^0(a), \\ f_{t,x}^{0*}(v) & \text{if } v \in \partial f_{t,x}^0([a,b]), \\ bv - f_{t,x}^0(b) & \text{if } v > \partial f_{t,x}^0(b), \end{cases}$$

and

$$\partial f_{t,x}^*(v) = \begin{cases} a & \text{if } v < \partial f_{t,x}^0(a), \\ \partial f_{t,x}^{0*}(v) & \text{if } v \in \partial f_{t,x}^0([a,b]), \\ b & \text{if } v > \partial f_{t,x}^0(b). \end{cases}$$

for  $\partial f_{t,x}^0([a,b]) := \bigcup_{u \in [a,b]} \partial f_{t,x}^0(u)$ . As we know that  $f^*$  is convex,  $\partial f^*$  is non-decreasing, and therefore, we have  $\partial f^* \subseteq [a,b]$ . As the subgradient of  $f_{t,x}^*$  is similarly bounded, we can obtain the same Lipschitz constant as Example 4.

#### C.5 Formal Statement and Proof of Lemma 3

The informal statement in Lemma 3 can be formally stated as follows:

**Lemma 8** *Let us consider solution  $\eta_{\text{CMC}}^* \in L^2(\mathcal{T} \times \mathcal{X}, p_{\text{obs}}) : \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}$  for some fixed policy. Define the orthogonal function class as the principal components obtained by applying the kernel PCA to the data so that  $\psi := \varphi^{\text{KPCA}} = (\varphi_1^{\text{KPCA}}, \dots, \varphi_D^{\text{KPCA}})$ . Suppose that kernel  $k(\cdot, \cdot)$  used by the kernel PCA is universal in  $L^2(\mathcal{T} \times \mathcal{X}, p_{\text{obs}})$  and has spectral decomposition  $k = \sum_{d=1}^D \lambda_d^* \varphi_d^* \otimes \bar{\varphi}_d^*$ , for  $\lambda_1^* \geq \lambda_2^* \geq \dots \geq 0$ . Let us introduce feature map  $\Phi : (t, x) \mapsto \sum_{d=1}^{\infty} \lambda_d^* \varphi_d^*(t, x) \varphi_d^*$  so that  $k((t, x), (t', x')) = \langle \Phi(t, x), \Phi(t', x') \rangle_{\mathcal{H}}$ . Now, assume that there exist constants  $M_1$  and  $M_2$  such that  $\|\Phi(T, X)\|_{\mathcal{H}}^2 < M_1$  and  $\|\Phi(T, X) \otimes \bar{\Phi}(T, X)\| < M_2$  almost surely, where the norm of product  $\Phi \otimes \bar{\Phi}$  is defined as the Hilbert-Schmidt norm. Additionally, define  $c_d := \langle \eta_{\text{CMC}}^*, \varphi_d^* \rangle$  so that  $\Pi_{\varphi^*} \eta_{\text{CMC}}^* = \sum_{d=1}^D c_d \varphi_d$*

and further assume that  $\lim_{D \rightarrow \infty} \left\{ D \left( \max_{1 \leq d \leq D} \frac{c_d^2}{\lambda_d^*} \right) \left( \sum_{d=D+1}^{\infty} \lambda_d^* \right) \right\} = 0$ . Then, we have

$$\lim_{D \rightarrow \infty} \text{plim}_{n \rightarrow \infty} \|\Pi_{\psi} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\| = 0,$$

where  $\text{plim}$  denotes convergence in probability.

**Proof** Due to the universal property of the kernel, we can approximate  $\eta_{\text{KCMC}}^*$  with the element of  $\mathcal{H}$  arbitrarily well. Therefore, for any  $\varepsilon > 0$ , we can take  $D$  such that  $\|\Pi_{\varphi^*} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\| \leq \frac{\varepsilon}{3}$  for  $\varphi^* := (\varphi_1^*, \dots, \varphi_D^*)^T$  and  $D \left( \max_{1 \leq d \leq D} \frac{c_d^2}{\lambda_d^*} \right) \left( \sum_{d=D+1}^{\infty} \lambda_d^* \right) \leq \frac{\varepsilon}{3}$ . For such a choice of  $D$ , we can prove the statement of the theorem by showing  $\text{plim}_{n \rightarrow \infty} \|\Pi_{\varphi^{\text{KPCA}}} \Pi_{\varphi^*} \eta_{\text{CMC}}^* - \Pi_{\varphi^*} \eta_{\text{CMC}}^*\| \leq \frac{\varepsilon}{3}$ . This is because

$$\begin{aligned} & \|\Pi_{\psi} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\| \\ &= \|\Pi_{\varphi^{\text{KPCA}}} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\| \\ &\leq \|\Pi_{\varphi^{\text{KPCA}}} (\eta_{\text{CMC}}^* - \Pi_{\varphi^*} \eta_{\text{CMC}}^*)\| + \|\Pi_{\varphi^{\text{KPCA}}} \Pi_{\varphi^*} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\| \\ &\leq \|\Pi_{\varphi^*} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\| + \|\Pi_{\varphi^{\text{KPCA}}} \Pi_{\varphi^*} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\| \\ &\leq \|\Pi_{\varphi^*} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\| + \|\Pi_{\varphi^{\text{KPCA}}} \Pi_{\varphi^*} \eta_{\text{CMC}}^* - \Pi_{\varphi^*} \eta_{\text{CMC}}^*\| + \|\Pi_{\varphi^*} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\| \\ &= \|\Pi_{\varphi^{\text{KPCA}}} \Pi_{\varphi^*} \eta_{\text{CMC}}^* - \Pi_{\varphi^*} \eta_{\text{CMC}}^*\| + 2\|\Pi_{\varphi^*} \eta_{\text{CMC}}^* - \eta_{\text{CMC}}^*\|, \end{aligned}$$

where we used the non-expansive property of the projection operator in the second line.

Now, as we know  $\Pi_{\varphi^*} \eta_{\text{CMC}}^* = \sum_{d=1}^D c_d \varphi_d^*$ , we have

$$\begin{aligned} \|\Pi_{\varphi^{\text{KPCA}}} \Pi_{\varphi^*} \eta_{\text{CMC}}^* - \Pi_{\varphi^*} \eta_{\text{CMC}}^*\|^2 &= \left\| \Pi_{\varphi^{\text{KPCA}}} \sum_{d=1}^D c_d \varphi_d^* - \sum_{d=1}^D c_d \varphi_d^* \right\|^2 \\ &= \left\| \sum_{d=1}^D \frac{c_d}{\sqrt{\lambda_d^*}} \cdot \sqrt{\lambda_d^*} (\Pi_{\varphi^{\text{KPCA}}} - 1) \varphi_d^* \right\|^2 \\ &\leq \max_{1 \leq d \leq D} \left( \frac{c_d^2}{\lambda_d^*} \right) \cdot \left\| \sum_{d=1}^D \sqrt{\lambda_d^*} (\Pi_{\varphi^{\text{KPCA}}} - 1) \varphi_d^* \right\|^2 \\ &\leq \max_{1 \leq d \leq D} \left( \frac{c_d^2}{\lambda_d^*} \right) \cdot D \sum_{d=1}^D \lambda_d^* \|(\Pi_{\varphi^{\text{KPCA}}} - 1) \varphi_d^*\|^2 \\ &\leq \max_{1 \leq d \leq D} \left( \frac{c_d^2}{\lambda_d^*} \right) \cdot D \sum_{d=1}^{\infty} \lambda_d^* \|(\Pi_{\varphi^{\text{KPCA}}} - 1) \varphi_d^*\|^2 \\ &= D \cdot \max_{1 \leq d \leq D} \left( \frac{c_d^2}{\lambda_d^*} \right) \cdot \sum_{d=1}^{\infty} \lambda_d^* (1 - \|\Pi_{\varphi^{\text{KPCA}}} \varphi_d^*\|^2). \end{aligned}$$

To show that the summation in the last line converges to zero as  $n \rightarrow \infty$ , we use the convergence result of the subspace learned by the kernel PCA in Blanchard et al. (2007, Theorem 3.1.). Their result states that under conditions  $\|\Phi(T, X)\|_{\mathcal{H}}^2 < M_1$  and  $\|\Phi(T, X) \otimes \bar{\Phi}(T, X)\| < M_2$  almost surely, we have

$$0 \leq \mathbb{E} \|\Pi_{\varphi^*} \Phi(T, X)\|^2 - \mathbb{E} \|\Pi_{\varphi^{\text{KPCA}}} \Phi(T, X)\|^2 \xrightarrow{P} 0$$

as  $n \rightarrow \infty$ , where  $\xrightarrow{p}$  indicate the convergence in probability.

Indeed, we can re-write this condition as

$$\begin{aligned}
 & \mathbb{E} \|\Pi_{\varphi^*} \Phi(T, X)\|^2 - \mathbb{E} \|\Pi_{\varphi^{\text{KPCA}}} \Phi(T, X)\|^2 \\
 &= \mathbb{E} \left\| \Pi_{\varphi^*} \sum_{d=1}^{\infty} \sqrt{\lambda_d^*} \varphi_d^*(T, X) \varphi_d^* \right\|^2 - \mathbb{E} \left\| \Pi_{\varphi^{\text{KPCA}}} \sum_{d=1}^{\infty} \sqrt{\lambda_d^*} \varphi_d^*(T, X) \varphi_d^* \right\|^2 \\
 &= \mathbb{E} \left\| \sum_{d=1}^{\infty} \sqrt{\lambda_d^*} \varphi_d^*(T, X) \Pi_{\varphi^*} \varphi_d^* \right\|^2 - \mathbb{E} \left\| \sum_{d=1}^{\infty} \sqrt{\lambda_d^*} \varphi_d^*(T, X) \Pi_{\varphi^{\text{KPCA}}} \varphi_d^* \right\|^2 \\
 &= \mathbb{E} \left\| \sum_{d=1}^D \sqrt{\lambda_d^*} \varphi_d^*(T, X) \varphi_d^* \right\|^2 - \mathbb{E} \left\| \sum_{d=1}^{\infty} \sqrt{\lambda_d^*} \varphi_d^*(T, X) \Pi_{\varphi^{\text{KPCA}}} \varphi_d^* \right\|^2 \\
 &= \sum_{d=1}^D \lambda_d^* \mathbb{E} |\varphi_d^*(T, X)|^2 \|\varphi_d^*\|^2 - \sum_{d=1}^{\infty} \sum_{k=1}^{\infty} \sqrt{\lambda_d^*} \sqrt{\lambda_k^*} \langle \Pi_{\varphi^{\text{KPCA}}} \varphi_d^*, \Pi_{\varphi^{\text{KPCA}}} \varphi_k^* \rangle_{\mathcal{H}} \mathbb{E} [\varphi_d^*(T, X), \varphi_k^*(T, X)] \\
 &= \sum_{d=1}^D \lambda_d^* \|\varphi_d^*\|^2 \cdot \|\varphi_d^*\|^2 - \sum_{d=1}^{\infty} \sum_{k=1}^{\infty} \sqrt{\lambda_d^*} \sqrt{\lambda_k^*} \langle \Pi_{\varphi^{\text{KPCA}}} \varphi_d^*, \Pi_{\varphi^{\text{KPCA}}} \varphi_k^* \rangle_{\mathcal{H}} \langle \varphi_d^*, \varphi_k^* \rangle_{\mathcal{H}} \\
 &= \sum_{d=1}^D \lambda_d^* - \sum_{d=1}^{\infty} \sum_{k=1}^{\infty} \sqrt{\lambda_d^*} \sqrt{\lambda_k^*} \langle \Pi_{\varphi^{\text{KPCA}}} \varphi_d^*, \Pi_{\varphi^{\text{KPCA}}} \varphi_k^* \rangle_{\mathcal{H}} \mathbb{1}_{d=k} \\
 &= \sum_{d=1}^D \lambda_d^* - \sum_{d=1}^{\infty} \lambda_d^* \|\Pi_{\varphi^{\text{KPCA}}} \varphi_d^*\|^2 \\
 &= \sum_{d=1}^{\infty} \lambda_d^* - \sum_{d=D+1}^{\infty} \lambda_d^* - \sum_{d=1}^{\infty} \lambda_d^* \|\Pi_{\varphi^{\text{KPCA}}} \varphi_d^*\|^2 \\
 &= \sum_{d=1}^{\infty} \lambda_d^* (1 - \|\Pi_{\varphi^{\text{KPCA}}} \varphi_d^*\|^2) - \sum_{d=D+1}^{\infty} \lambda_d^* \\
 &\xrightarrow{p} 0.
 \end{aligned}$$

Therefore, we have

$$\begin{aligned}
 & \text{plim}_{n \rightarrow \infty} \|\Pi_{\varphi^{\text{KPCA}}} \Pi_{\varphi^*} \eta_{\text{CMC}}^* - \Pi_{\varphi^*} \eta_{\text{CMC}}^*\|^2 \\
 &\leq \text{plim}_{n \rightarrow \infty} D \cdot \max_{1 \leq d \leq D} \left( \frac{c_d^2}{\lambda_d^*} \right) \cdot \left( \mathbb{E} \|\Pi_{\varphi^*} \Phi(T, X)\|^2 - \mathbb{E} \|\Pi_{\varphi^{\text{KPCA}}} \Phi(T, X)\|^2 + \sum_{d=D+1}^{\infty} \lambda_d^* \right) \\
 &\leq D \cdot \max_{1 \leq d \leq D} \left( \frac{c_d^2}{\lambda_d^*} \right) \cdot \text{plim}_{n \rightarrow \infty} \left( \mathbb{E} \|\Pi_{\varphi^*} \Phi(T, X)\|^2 - \mathbb{E} \|\Pi_{\varphi^{\text{KPCA}}} \Phi(T, X)\|^2 \right) + \frac{\varepsilon}{3} \\
 &= \frac{\varepsilon}{3},
 \end{aligned}$$

which concludes the proof. ■

This lemma indicates that convergence of the specification bias to zero can be guaranteed by employing a universal kernel (Micchelli et al., 2006). However, the current proof lacks the convergence rate, which is an open question for future work.

In the above proof, the most obscure assumption for readers would be the condition that  $\lim_{D \rightarrow \infty} \left\{ D \left( \max_{1 \leq d \leq D} \frac{c_d^2}{\lambda_d^*} \right) \left( \sum_{d=D+1}^{\infty} \lambda_d^* \right) \right\} = 0$ . This assumption can be approximately decomposed into two slightly more interpretable assumptions. First, it requires that  $\max_{1 \leq d \leq D} \frac{c_d^2}{\lambda_d^*}$  is bounded so that basis function  $\varphi_d^*$  for smaller spectrum  $\lambda_d^*$  constitutes proportionally smaller part of function  $\eta_{\text{CMC}}^*$ . For example, if  $\eta_{\text{CMC}}^*$  happens to coincide with a function generated from the Gaussian process with kernel  $k(\cdot, \cdot)$ , it will satisfy this property with high probability. Second, it requires that  $\sum_{d=D+1}^{\infty} \lambda_d^* = o_p(\frac{1}{D})$  so that the decay of spectrum  $\lambda_d^*$  is fast. In fact, this assumption is similar to another assumption  $\|\Phi(T, X)\|_{\mathcal{H}}^2 \leq M_1$  we made, because this assumption leads to  $M_1 \geq \mathbb{E}\|\Phi(T, X)\|_{\mathcal{H}}^2 = \mathbb{E} \left\| \sum_{d=1}^{\infty} \sqrt{\lambda_d^*} \varphi_d(T, X) \varphi_d \right\|^2 = \sum_{d=1}^{\infty} \lambda_d^* \mathbb{E} |\varphi_d(T, X)|^2 \cdot \|\varphi_d\|^2 = \sum_{d=1}^{\infty} \lambda_d^*$ , which implies convergence of its tail part  $\sum_{d=D+1}^{\infty} \lambda_d^* \rightarrow 0$ .

### C.6 Details of Example 6

**Proof** We know the analytical form of the dual solution,  $\eta_{\text{CMC}}^*(t, x) = \left( \frac{\pi(t|x)}{p_{\text{obs}}(t|x)} \right) Q(t, x)$  as discussed earlier. Therefore, we can take  $D = 1$  and set  $\psi_1(t, y) = \left( \frac{\pi(t|x)}{p_{\text{obs}}(t|x)} \right) Q(t, x)$  to meet condition (28) in Theorem 2 to obtain the kernel conditional moment constraint with no specification error. Therefore, we can solve (15) in the case of the box constraints as

$$V_{\text{inf}}^{\text{CMC}}(\pi) = \min_{a_w(t,x) \leq w(y,t,x) \leq b_w(t,x)} \mathbb{E}_{\text{obs}}[w(Y, T, X) \pi(T|X) Y]$$

subject to

$$\mathbb{E}_{\text{obs}}[w(Y, T, X) \pi(T|X) Q(T, X)] = \mathbb{E}_{\text{obs}} \left[ \left( \frac{\pi(T|X)}{p_{\text{obs}}(T|X)} \right) Q(T, X) \right],$$

where  $Q(t, x)$  denotes the  $\tau(t, x)$ -quantile of the conditional distribution of  $Y$  given  $T = t$  and  $X = x$  for  $\tau(t, x) := \frac{1/p_{\text{obs}}(t|x) - a_w(t, x)}{b_w(t, x) - a_w(t, x)}$ .

In the case of marginal sensitivity model (8) by Tan (2006), the expression for  $\tau(t, x)$  can be simplified as  $\tau(x) = \frac{1}{1+\Gamma}$ , and therefore,  $Q(t, x)$  can be estimated from finite samples by the standard quantile regression. The quantile balancing (QB) estimator introduced by Dorn and Guo (2022) relies on this property and estimates the sharp lower bound by solving the empirical version of the above problem with quantile estimate  $\hat{Q}(t, x)$  obtained by the quantile regression.

In addition to showing that their two-stage estimator can be considered a special case of the KCMC estimator, we can also argue that the KCMC estimator can be sharper than the quantile balancing estimator. For that purpose, we pick an equivalent choice of feature sets for the KCMC estimator and the QB estimator. When function set  $\{\psi_d(t, x)\}_{d=1}^D$  is used by the KCMC estimator, we can choose  $\left\{ \left( \frac{p_{\text{obs}}(t|x)}{\pi(t|x)} \right) \psi_d(t, x) \right\}_{d=1}^D$  as the feature of the linear quantile regression. Then, when the solution of the quantile regression is

$\hat{Q}(t, x) = \hat{\eta}_{\text{KCMC}}^T \psi(t, x)$ , the QB estimator becomes as tight as the KCMC estimator. However, in general, the solution of the quantile regression may not coincide  $\hat{\eta}_{\text{KCMC}}$  even at an infinite sample limit, since the objective of the quantile regression is not equivalent to the dual objective of the KCMC estimator, and in such cases, the KCMC estimator becomes tighter than the QB estimator.

As our estimator generalizes the previous work, our estimator overcomes some drawbacks of the quantile balancing estimators. As discussed in Section 1, the quantile balancing estimator cannot handle policy learning and the f-constraint. Policy learning is also difficult with the quantile balancing estimator because taking the derivative with respect to policy requires differentiability of the solution of the above linear programming with respect to parameter  $Q$ . Moreover, the quantile balancing method is designed only for the box constraints and does not have a proper extension to the f-sensitivity model (10). In contrast, our estimator of sharper bound  $V_{\text{inf}}^{\text{CMC}}$  based on the kernel method can naturally handle the above-mentioned generalized cases of sensitivity analysis. ■

### C.7 Proof and full version of Theorem 4

Let us introduce loss function  $\ell : \Theta \times \mathcal{Z} \rightarrow \mathbb{R}$ , where  $\Theta \subseteq \mathbb{R}^K$  for some  $K$  and  $\mathcal{Z} := \mathcal{Y} \times \mathcal{T} \times \mathcal{X}$ . To prove the consistency of policy evaluation, we assume the following set of regularity conditions for this loss function:

#### Condition 1 (Regularity of loss function I)

1.  $\theta \mapsto \ell_\theta(z)$  is continuous for any  $z \in \mathcal{Z}$ .
2.  $\mathbb{E}|\ell_\theta(Z)| < \infty$  for any  $\theta \in \Theta$ .
3.  $\theta^* \in \arg \min_{\theta \in \Theta} \mathbb{E}[\ell_\theta(Z)]$  is unique.
4.  $\theta^*$  is well-separated, i.e.,  $\inf_{\theta: \|\theta - \theta^*\| > \varepsilon} \mathbb{E}[\ell_\theta(Z)] > \mathbb{E}[\ell_{\theta^*}(Z)]$  for any  $\varepsilon > 0$ .
5.  $\mathbb{E}[G_\varepsilon(Z)] < \infty$  for  $G_\varepsilon(z) := \sup_{\theta \in \Theta: \|\theta - \theta^*\| \leq \varepsilon} |\ell_\theta(z)|$  for some  $\varepsilon > 0$ .

Additionally, we will need the following convergence lemmas:

#### Lemma 9 (Uniform convergence on compact space (Van de Geer, 2020, Lemma 7.2.1.))

Assume that parameter space  $(\Theta, \|\cdot\|)$  is compact and satisfies Condition 1. Then,

$$\sup_{\theta \in \Theta} \left| \hat{\mathbb{E}}_n[\ell_\theta(Z)] - \mathbb{E}[\ell_\theta(Z)] \right| \xrightarrow{P} 0.$$

#### Lemma 10 (Consistency of convex M-estimation (Van de Geer, 2020, Lemma 7.2.2.))

Suppose  $\theta \mapsto \ell_\theta(z)$  is convex for any  $z \in \mathcal{Z}$  and that  $\Theta \subseteq \mathbb{R}^k$  is convex. Then, for M-estimator  $\hat{\theta}_n \in \arg \min_{\theta \in \Theta} \hat{\mathbb{E}}_n[\ell_\theta(Z)]$ , we have  $\hat{\theta}_n \xrightarrow{P} \theta^*$ .

With these lemmas, we can prove Theorem 4.

**Theorem 3 (Consistency of policy evaluation (full version))** Define the parameter space of  $(\eta_f, \eta_{\text{KCMC}})$  as  $\Theta \subseteq \mathbb{R}_+ \times \mathbb{R}^D$ . Further, define  $\theta^* := (\eta_f^*, \eta_{\text{KCMC}}^*)$  as the solution of dual problem (26) for  $V_{\text{inf}}^{\text{KCMC}}(\pi)$  and  $\hat{\theta}_n := (\hat{\eta}_f, \hat{\eta}_{\text{KCMC}})$  as the solution to dual problem (27) for  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi)$ . Define  $\ell : \Theta \times \mathcal{Z} \rightarrow \mathbb{R}$  as

$$\ell_\theta(t, y, x) := \eta_f \gamma - \eta_{\text{KCMC}}^T \boldsymbol{\psi}(t, x) + \eta_f f_{t,x}^* \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi}(t, x) - r(y, t, x)}{\eta_f} \right)$$

so that it is the negative version of dual objectives (26) and (27) before taking the expectations. Now, assume Condition 1 holds for the above loss function. Then, we have  $\hat{\theta}_n \xrightarrow{P} \theta^*$  and  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi) \xrightarrow{P} V_{\text{inf}}^{\text{KCMC}}(\pi)$ .

### Proof

As our dual problem for policy evaluation (26) and (27) are concave maximization, we can immediately apply the above lemma as follows. We can immediately apply Lemma 10 and get  $\hat{\theta}_n \xrightarrow{P} \theta^*$ . Thus,  $\hat{\theta}_n$  tend to the inside of compact set  $\{\theta \in \Theta : \|\theta - \theta^*\| \leq \varepsilon\}$ , in which we have the uniform convergence of  $\hat{\mathbb{E}}_n[\ell_\theta(Z)]$  to  $\mathbb{E}[\ell_\theta(Z)]$  by Lemma 9. Therefore, we have  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi) = -\hat{\mathbb{E}}_n[\ell_{\hat{\theta}_n}(Z)] \xrightarrow{P} -\mathbb{E}[\ell_{\hat{\theta}_n}(Z)] \xrightarrow{P} -\mathbb{E}[\ell_{\theta^*}(Z)] = V_{\text{inf}}^{\text{KCMC}}(\pi)$ .  $\blacksquare$

In practice, it is difficult to check some of the conditions in 1, such as the integrability assumption  $\mathbb{E}|\ell_\theta| < \infty$  for any  $\theta \in \Theta$  as well as the uniqueness of the solution. However, it is possible in some cases to verify  $L^1$  envelope condition  $\mathbb{E}[G_\varepsilon] < \infty$ , because local Lipschitzness of  $f_{t,x}^* : \Theta \rightarrow \mathbb{R}$  implies the existence of such  $\varepsilon$ . For example, for the box constraints of Example 3, we know that  $f_{t,x}^*$  is upper bounded by  $b_{\bar{w}}(t, x)$ . For f-constraints (12), the conjugate function  $f^*$  for many choices of f-divergence (such as Kullback-Leibler (KL), squared Hellinger, etc.) is locally Lipschitz.

## C.8 Proof and full version of Theorem 5

Here, we simply utilize Lemma 10 analogously to Theorem 4 to prove the theorem.

**Theorem 4 (Consistency of concave policy learning (full version))** Assume concave policy class  $\{\pi_\beta(t|x) : \beta \in \mathcal{B}\}$  with convex parameter space  $\mathcal{B}$  satisfying that  $\beta \mapsto \pi_\beta(t|x)y$  is concave for any  $y \in \mathcal{Y}$ ,  $t \in \mathcal{T}$  and  $x \in \mathcal{X}$ . Assume Condition 1 holds for loss function (29). Then, we have  $\hat{\theta}_n \xrightarrow{P} \theta^*$  and  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) \xrightarrow{P} V_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*})$ .

**Proof** Due to the concavity of policy class  $\{\pi_\beta(t|x) : \beta \in \mathcal{B}\}$ , we know that  $\beta \mapsto \mathbb{E} \left[ \tilde{w}(Y, T, X) \left( \frac{\pi_\beta(T|X)}{p_{\text{obs}}(T|X)} \right) Y \right]$  is concave for any  $t \in \mathcal{T}$  and  $x \in \mathcal{X}$ . Then, we can see that

$$\begin{aligned} \max_{\beta \in \mathcal{B}} V_{\text{inf}}^{\text{KCMC}} &= \max_{\beta \in \mathcal{B}} \max_{\substack{\eta_{\text{KCMC}} \in \mathbb{R}^D \\ \eta_f > 0}} \min_{\tilde{w}} \mathbb{E} \left[ \tilde{w}(Y, T, X) \left( \frac{\pi_\beta(T|X)}{p_{\text{obs}}(T|X)} \right) Y \right] \\ &\quad - \mathbb{E} [(\tilde{w} - 1) \eta_{\text{KCMC}}^T \boldsymbol{\psi}] + \eta_f (\mathbb{E}[f_{T,X}(\tilde{w})] - \gamma) \\ &= \max_{\beta \in \mathcal{B}} \max_{\substack{\eta_{\text{KCMC}} \in \mathbb{R}^D \\ \eta_f > 0}} \mathbb{E}[-\ell_\theta(Y, T, X)] \end{aligned}$$

$$= \max_{\theta \in \Theta} \mathbb{E}[-\ell_\theta(Y, T, X)]$$

is a concave maximization problem, because  $\mathbb{E}[\ell_\theta]$  is the pointwise infimum of concave functions. Thus, we can apply Lemma 10 and get  $\hat{\theta}_n \xrightarrow{P} \theta^*$ , which implies  $\hat{\theta}_n$  tend to the inside of compact set  $\{\theta \in \Theta : \|\theta - \theta^*\| \leq \varepsilon\}$ , where we have uniform convergence guarantee of  $\hat{\mathbb{E}}_n[\ell_\theta(Z)]$  due to Lemma 9. Therefore, we have

$$\begin{aligned} 0 &\leq \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) - \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*}) \\ &= -\hat{\mathbb{E}}_n[\ell_{\hat{\theta}_n}(Z)] + \hat{\mathbb{E}}_n[\ell_{\theta^*}(Z)] \\ &= \left[ -\hat{\mathbb{E}}_n[\ell_{\hat{\theta}_n}(Z)] + \mathbb{E}[\ell_{\hat{\theta}_n}(Z)] \right] + \left[ \hat{\mathbb{E}}_n[\ell_{\theta^*}(Z)] - \mathbb{E}[\ell_{\theta^*}(Z)] \right] - \left[ \mathbb{E}[\ell_{\hat{\theta}_n}(Z)] - \mathbb{E}[\ell_{\theta^*}(Z)] \right] \\ &\leq \left[ -\hat{\mathbb{E}}_n[\ell_{\hat{\theta}_n}(Z)] + \mathbb{E}[\ell_{\hat{\theta}_n}(Z)] \right] + \left[ \hat{\mathbb{E}}_n[\ell_{\theta^*}(Z)] - \mathbb{E}[\ell_{\theta^*}(Z)] \right] \\ &\xrightarrow{P} 0, \end{aligned}$$

which implies  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) \xrightarrow{P} \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*})$ . Lastly, due to the law of large numbers, we have  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*}) \xrightarrow{P} V_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*})$ , which concludes the proof.  $\blacksquare$

### C.9 Proof and full version of Theorem 6

Let us define  $\Theta_\varepsilon := \{(\eta, \beta) : \|\eta - \eta^*(\beta)\| \leq \varepsilon, \beta \in \mathcal{B}\}$  for  $\eta^*(\beta) := \arg \min_{\eta \in H} \mathbb{E}[\ell_{\beta, \eta}(Z)]$ , which is the  $\varepsilon$ -neighborhood of the partially optimized parameters. Here, we need the following regularity conditions on the loss function  $\ell_\theta(y, t, x) = \ell_{\eta, \beta}(y, t, x)$ .

#### Condition 2 (Regularity of loss function II)

1.  $\theta \mapsto \ell_\theta(z)$  is continuous for any  $z \in \mathcal{Z}$ .
2.  $\mathbb{E}|\ell_\theta(Z)| < \infty$  for any  $\theta \in \Theta$ .
3.  $\eta^*(\beta) \in \arg \min_{\eta \in H} \mathbb{E}[\ell_{\eta, \beta}(Z)]$  is unique for any  $\beta \in \mathcal{B}$ .
4.  $\eta^*(\beta)$  is uniformly well-separated, i.e., for any  $\varepsilon > 0$ ,

$$\inf_{\theta \in \Theta \setminus \Theta_\varepsilon} \left| \mathbb{E}[\ell_{\eta, \beta}(Z)] - \mathbb{E}[\ell_{\eta^*(\beta), \beta}(Z)] \right| > 0.$$

5.  $f_{t,x}^*$  is uniformly Lipschitz continuous, i.e., there exists  $L > 0$  such that

$$\sup_{x \in \mathcal{X}, t \in \mathcal{T}, v \in \mathbb{R}} \sup_{\delta f^* \in \partial f_{t,x}^*} |\delta f^*(v)| \leq L.$$

6.  $\eta^*(\beta)$  is uniformly bounded so that 1) there exist  $\underline{\eta}_f, \bar{\eta}_f \in \mathbb{R}$  such that  $0 < \underline{\eta}_f \leq \eta_f^*(\beta) \leq \bar{\eta}_f < \infty$  for any  $\beta \in \mathcal{B}$ , and 2)  $\bar{\eta}_{\text{KCMC}} := \sup_{\beta \in \mathcal{B}} \|\eta_{\text{KCMC}}^*(\beta)\| < \infty$ .

7.  $\mathbb{E}[G^\psi(Z)] < \infty$  and  $\mathbb{E}[G_\varepsilon^{f^*}(Z)] < \infty$  for

$$G^\psi(z) := \|\psi(t, x)\|,$$

$$G_\varepsilon^{f^*}(z) := \sup_{\theta \in \Theta_\varepsilon} \left| f_{t,x}^* \left( \frac{\eta_{\text{KCMC}}^T \psi(t, x) - \left( \frac{\pi_\beta(t|x)}{p_{\text{obs}}(t|x)} \right) y}{\eta_f} \right) \right|$$

for some  $\varepsilon > 0$ .

8.  $\mathbb{E}[R(Z)] < \infty$  for  $R(z) := \sup_{\beta \in \mathcal{B}} \left| \left( \frac{\pi_\beta(t|x)}{p_{\text{obs}}(t|x)} \right) y \right|$ .

**Remark 3** In Condition 2, we required uniform well-separatedness assumption and uniform Lipschitzness of  $v \mapsto f_{t,x}^*(v)$ , which may not seem obvious to the readers. To put the former condition in slightly more practical terms, we can consider the Hessian of  $\eta \mapsto \mathbb{E}\ell_\theta(Z)$ . If the smallest eigenvalues of the Hessian of  $\mathbb{E}\ell_\theta(Z)$  is uniformly lower bounded so that there exists  $\delta > 0$  such that  $\nabla_\eta \nabla_\eta \mathbb{E}\ell_\theta(Z) - \delta I$  is positive semidefinite for any  $\theta \in \Theta_\varepsilon$ , we know that  $\mathbb{E}\ell_{\eta,\beta}(z) - \mathbb{E}\ell_{\eta^*(\beta),\beta}(z) \geq \frac{1}{2}\delta \|\eta - \eta^*(\beta)\|^2$  for any  $(\eta, \beta) \in \Theta_\varepsilon$ . With regard to the latter condition of uniform Lipschitzness, we can think of Example 5. Since the subgradient of  $f^*$  can be uniformly bounded by  $b_{\bar{w}}(t, x)$ , we require that  $b_{\bar{w}}(t, x)$  be uniformly bounded. In the case of Tan's box constraint (8), we know that  $b_{\bar{w}}(t, x) = p_{\text{obs}}(t|x) - \Gamma(1 - p_{\text{obs}}(t|x))$  and  $p_{\text{obs}}(t|x) \leq 1$ , which implies the uniform Lipschitz condition is always satisfied.

We begin the proof by showing uniform convergence  $\sup_{\theta \in \Theta_\varepsilon} \left| \hat{\mathbb{E}}_n \ell_\theta(Z) - \mathbb{E}\ell_\theta(Z) \right| \xrightarrow{P} 0$  for some  $\varepsilon > 0$ . For the proof of Theorem 6, we need the following lemmas:

**Lemma 13 (Stability of VC functions (Vaart and Wellner, 1996, Lemma 2.6.18))**

Let  $\mathcal{F}, \mathcal{G}$  be VC classes of functions on  $\mathcal{Z}$  and  $\mathbb{R}^D$ . Let  $h_1 : \mathcal{Z} \rightarrow \mathbb{R}$ ,  $h_2 : \mathcal{Z} \rightarrow \mathbb{R}^D$  be fixed functions. Then, product functions  $h_1 \cdot \mathcal{F} := \{h_1 \cdot f : f \in \mathcal{F}\}$  and composite functions  $\mathcal{G} \circ h_2 := \{g \circ h_2 : g \in \mathcal{G}\}$  are VC.

**Definition 14 (Covering number (Van de Geer, 2020, Definition 6.1.2.))** A class of functions  $\mathcal{G}^{(\delta)}$  is said to be a  $\delta$ -covering of  $\mathcal{G}$  on  $\mathcal{Z}$  with respect to  $L_1(\mathcal{Z}, \mathbb{E})$  norm if

$$\sup_{g \in \mathcal{G}} \inf_{\tilde{g} \in \mathcal{G}^{(\delta)}} \mathbb{E} |g(Z) - \tilde{g}(Z)| \leq \delta.$$

Additionally, the covering number of  $\mathcal{G}$  is defined as the minimal cardinality of its  $\delta$ -covering so that

$$N_1(\delta, \mathcal{G}, \mathbb{E}) = \min \left\{ \left| \mathcal{G}^{(\delta)} \right| : \mathcal{G}^{(\delta)} \text{ is a } \delta\text{-covering of } \mathcal{G} \text{ w.r.t. } L_1(\mathcal{Z}, \mathbb{E}) \text{ norm} \right\}.$$

**Lemma 15 (Covering number of a VC class (Van de Geer, 2020, Theorem 6.4.1))**

Let  $Q$  be any probability measure on  $\mathcal{Z}$  and let  $N_1(\cdot, \mathcal{G}, \mathbb{E}_Q)$  be the covering number of  $\mathcal{G}$ . For a VC class  $\mathcal{G}$  with VC dimension  $V$  and bounded envelope  $G(z) := \sup_{g \in \mathcal{G}} |g(z)|$  such that  $\mathbb{E}_Q[G(Z)] < \infty$ , we have a constant  $A$  depending only on  $V$  (not on  $Q$ ) satisfying  $N_1(\delta \mathbb{E}_Q[G], \mathcal{G}, \mathbb{E}_Q) \leq \max(A\delta^{-2V}, \exp \delta/4)$  for any  $\delta > 0$ .

**Lemma 16 (Uniform convergence (Van de Geer, 2020, Theorem 6.1.2))** *Suppose  $\mathbb{E}[G(Z)] < \infty$  for  $G(z) := \sup_{g \in \mathcal{G}} g(z)$  and  $\frac{\log N_1(\delta, \mathcal{G}, \hat{\mathbb{E}}_n)}{n} \xrightarrow{P} 0$  for any  $\delta > 0$ . Then,  $\sup_{g \in \mathcal{G}} |\hat{\mathbb{E}}_n g - \mathbb{E}g| \xrightarrow{P} 0$ .*

**Lemma 17 ( $L_1(\mathcal{Z}, \mathbb{E})$  envelope of  $\eta \mapsto \partial_\eta \ell_{\eta, \beta}(z)$ )** *Under Condition 2, there exist  $\varepsilon > 0$  such that envelopes  $G_\varepsilon^{\partial_{\eta_f} \ell}(z) := \sup_{\delta \ell \in \partial_{\eta_f} \ell_\theta(z)} |\delta \ell|$  and  $G_\varepsilon^{\partial_{\eta_{\text{KCMC}}} \ell}(z) := \sup_{\delta \ell \in \partial_{\eta_{\text{KCMC}}} \ell_\theta(z)} \|\delta \ell\|$  satisfy  $\mathbb{E}G_\varepsilon^{\partial_{\eta_f} \ell}(Z) < \infty$  and  $\mathbb{E}G_\varepsilon^{\partial_{\eta_{\text{KCMC}}} \ell}(Z) < \infty$ .*

**Proof** Let us take  $\varepsilon > 0$  sufficiently small so that envelopes  $G^\psi(z)$ ,  $G_\varepsilon^{f^*}(z)$ , and  $R(z)$  are in  $L_1(\mathcal{Z}, \mathbb{E})$ . We can show that  $\{z \mapsto \partial_{\eta_f} \ell_\theta(z) : \theta \in \Theta_\varepsilon\}$  has  $L_1(\mathcal{Z}, \mathbb{E})$  envelope  $G_\varepsilon^{\partial_{\eta_f} \ell}(z) := \sup_{\delta_{\eta_f} \ell \in \partial_{\eta_f} \ell_\theta(z)} |\delta_{\eta_f} \ell| \leq \gamma + G_\varepsilon^{f^*}(z) + L \left( \frac{\bar{\eta}_{\text{KCMC}} G^\psi(z) + R(z)}{\eta_f - \varepsilon} \right)$  because

$$\begin{aligned} \partial_{\eta_f} \ell_\theta(z) &= \gamma + f_{t,x}^* \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi}(t, x) - \left( \frac{\pi_\beta(t|x)}{p_{\text{obs}}(t|x)} \right) y}{\eta_f} \right) \\ &\quad + \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi}(t, x) - \left( \frac{\pi_\beta(t|x)}{p_{\text{obs}}(t|x)} \right) y}{\eta_f} \right) \partial_{\eta_f} f_{t,x}^* \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi}(t, x) - \left( \frac{\pi_\beta(t|x)}{p_{\text{obs}}(t|x)} \right) y}{\eta_f} \right). \end{aligned}$$

Similarly, we can see  $\{z \mapsto \partial_{\eta_{\text{KCMC}}} \ell_\theta(z) : \theta \in \Theta_\varepsilon\}$  also has  $L_1(\mathcal{Z}, \mathbb{E})$  envelope  $G_\varepsilon^{\partial_{\eta_{\text{KCMC}}} \ell}(z) := \sup_{\delta \ell \in \partial_{\eta_{\text{KCMC}}} \ell_\theta(z)} \|\delta \ell\| \leq (L+1)G^\psi(z)$  as

$$\partial_{\eta_{\text{KCMC}}} \ell_\theta(z) = \boldsymbol{\psi}(t, x) \left\{ 1 - \partial f_{t,x}^* \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi}(t, x) - \left( \frac{\pi_\beta(t|x)}{p_{\text{obs}}(t|x)} \right) y}{\eta_f} \right) \right\},$$

which concludes the proof.  $\blacksquare$

**Lemma 18 (Uniform convergence over  $\Theta_\varepsilon$ )** *Assume policy class  $\{\pi_\beta(t|x) : \beta \in \mathcal{B}\}$  is VC and that Condition 2 is met. Then, there exists  $\varepsilon > 0$  such that*

$$\sup_{\theta \in \Theta_\varepsilon} \left| \hat{\mathbb{E}}_n \ell_\theta(Z) - \mathbb{E} \ell_\theta(Z) \right| \xrightarrow{P} 0.$$

**Proof** Remember that loss function  $\ell : \Theta \times \mathcal{Z} \rightarrow \mathbb{R}$  is defined as

$$\ell_\theta(t, y, x) := \eta_f \gamma - \eta_{\text{KCMC}}^T \boldsymbol{\psi}(t, x) + \eta_f f_{t,x}^* \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi}(t, x) - \left( \frac{\pi_\beta(t|x)}{p_{\text{obs}}(t|x)} \right) y}{\eta_f} \right).$$

We choose  $\varepsilon$  such that envelopes  $G^\psi(z)$ ,  $G_\varepsilon^{f^*}(z)$ , and  $R(z)$  are in  $L_1(\mathcal{Z}, \mathbb{E})$ . Then,  $\{z \mapsto \ell_\theta(z) : \theta \in \Theta_\varepsilon\}$  has  $L_1(\mathcal{Z}, \mathbb{E})$  envelope  $G_\varepsilon^\ell(z) := \sup_{\theta \in \Theta_\varepsilon} |\ell_\theta(z)| \leq \gamma \eta_f + \bar{\eta}_{\text{KCMC}} G_\varepsilon^\psi(z) + \eta_f G_\varepsilon^{f^*}(z)$ .

Now, we would like to show that the class of loss functions  $\mathcal{G}_\varepsilon := \{\ell_\theta : \theta \in \Theta_\varepsilon\}$  has  $\delta$ -covering satisfying  $\frac{N_1(\delta, \mathcal{G}_\varepsilon, \hat{\mathbb{E}}_n)}{n} \xrightarrow{p} 0$  for some  $\varepsilon > 0$ . First, we know that

$$\mathcal{G}'_\varepsilon = \{(y, t, x) \mapsto \eta_{\text{KCMC}}^T \psi(t, x) : (\eta, \beta) \in \Theta_\varepsilon\}$$

and

$$\mathcal{G}''_\varepsilon = \left\{ (y, t, x) \mapsto \left( \frac{\pi_\beta(t|x)}{p_{\text{obs}}(t|x)} \right) y : \beta \in \mathcal{B} \right\}$$

are VC classes by Lemma 13. Thus, by Lemma 15, there exist  $\frac{\delta}{3(L+1)}$ -covering of these satisfying  $\frac{\log N_1(\delta/3(L+1), \mathcal{G}'_\varepsilon, \hat{\mathbb{E}}_n)}{n} \xrightarrow{p} 0$  and  $\frac{\log N_1(\delta/3(L+1), \mathcal{G}''_\varepsilon, \hat{\mathbb{E}}_n)}{n} \xrightarrow{p} 0$ . We let  $H_{\text{KCMC}}^{(\delta)} = \{\eta_{\text{KCMC}, k}^{(\delta)} : k = 1, \dots, N_1(\delta/3(L+1), \mathcal{G}'_\varepsilon, \hat{\mathbb{E}}_n)\}$  and  $\mathcal{B}^{(\delta)} = \{\beta_k^{(\delta)} : k = 1, \dots, N_1(\delta/3(L+1), \mathcal{G}''_\varepsilon, \hat{\mathbb{E}}_n)\}$  denote the parameter sets corresponding to such coverings. Also, we can take  $\frac{\delta}{3M}$ -mesh of  $\{\eta_f : \eta_f \in [\eta_f - \varepsilon, \bar{\eta}_f + \varepsilon]\}$  as  $H_f^{(\delta)} := \{\frac{\delta k}{3M} : k \in \mathbb{Z}\} \cap [\eta_f - \varepsilon, \bar{\eta}_f + \varepsilon]$  for  $M := 2\mathbb{E}G_\varepsilon^{\partial\eta_f \ell}(Z)$ , where  $G_\varepsilon^{\partial\eta_f \ell}(Z)$  is as defined in Lemma 17. Clearly, the cardinality of  $H_f^{(\delta)}$  does not depend on the sample size  $n$  and  $\frac{\log |H_f^{(\delta)}|}{n} \rightarrow 0$ .

Now, we can show that  $\Theta^{(\delta)} := H_f^{(\delta)} \times H_{\text{KCMC}}^{(\delta)} \times \mathcal{B}^{(\delta)}$  becomes  $\delta$ -covering of  $\mathcal{G}_\varepsilon$ . For any  $(\eta_f, \eta_{\text{KCMC}}, \beta) \in \Theta_\varepsilon$ , we can take  $(\eta_f^{(\delta)}, \eta_{\text{KCMC}}^{(\delta)}, \beta^{(\delta)}) \in H_f^{(\delta)} \times H_{\text{KCMC}}^{(\delta)} \times \mathcal{B}^{(\delta)}$  such that

$$|\eta_f - \eta_f^{(\delta)}| \leq \frac{\delta}{3M},$$

$$\hat{\mathbb{E}}_n \left| \eta_{\text{KCMC}}^T \psi(T, X) - \eta_{\text{KCMC}}^{(\delta) T} \psi(T, X) \right| \leq \frac{\delta}{3(L+1)},$$

and

$$\hat{\mathbb{E}}_n \left| \left( \frac{\pi_\beta(T|X)}{p_{\text{obs}}(T|X)} \right) Y - \left( \frac{\pi_{\beta^{(\delta)}}(T|X)}{p_{\text{obs}}(T|X)} \right) Y \right| \leq \frac{\delta}{3(L+1)}.$$

Thus, we have

$$\begin{aligned} \hat{\mathbb{E}}_n \left| \ell_{\eta_f, \eta_{\text{KCMC}}, \beta}(Z) - \ell_{\eta_f^{(\delta)}, \eta_{\text{KCMC}}, \beta}(Z) \right| &\leq \hat{\mathbb{E}}_n \sup_{\substack{\delta_{\eta_f} \ell \in \partial_{\eta_f} \ell_\theta(Z) \\ \theta \in \Theta_\varepsilon}} |\delta_{\eta_f} \ell| |\eta_f - \eta_f^{(\delta)}| \\ &\leq \hat{\mathbb{E}}_n G_\varepsilon^{\partial\eta_f \ell}(Z) |\eta_f - \eta_f^{(\delta)}| \\ &\xrightarrow{p} \frac{M}{2} \cdot \frac{\delta}{3M} \\ &< \frac{\delta}{3}. \end{aligned} \tag{31}$$

and

$$\begin{aligned} &\hat{\mathbb{E}}_n \left| \ell_{\eta_f^{(\delta)}, \eta_{\text{KCMC}}, \beta}(Z) - \ell_{\eta_f^{(\delta)}, \eta_{\text{KCMC}}, \beta^{(\delta)}}(Z) \right| \\ &\leq \hat{\mathbb{E}}_n \left| J_{\eta_f^{(\delta)}} \left( \eta_{\text{KCMC}}^T \psi(T, X), \left( \frac{\pi_\beta(T|X)}{p_{\text{obs}}(T|X)} \right) Y, T, X \right) - J_{\eta_f^{(\delta)}} \left( \eta_{\text{KCMC}}^{(\delta) T} \psi(T, X), \left( \frac{\pi_{\beta^{(\delta)}}(T|X)}{p_{\text{obs}}(T|X)} \right) Y, T, X \right) \right| \end{aligned}$$

$$\begin{aligned}
 &\leq \hat{\mathbb{E}}_n \left[ \sup_{\substack{\delta J \in \partial(h,r) J_{\eta_f^{(\delta)}}(h,r,t,x), \\ h,r \in \mathbb{R}, t \in \mathcal{T}, x \in \mathcal{X}}} \|\delta J\|_\infty \cdot \left\| \begin{pmatrix} \eta_{\text{KCMC}}^T \boldsymbol{\psi}(T, X) - \eta_{\text{KCMC}}^{(\delta)T} \boldsymbol{\psi}(T, X) \\ \left( \frac{\pi_\beta(T|X)}{p_{\text{obs}}(T|X)} \right) Y - \left( \frac{\pi_{\beta^{(\delta)}}(T|X)}{p_{\text{obs}}(T|X)} \right) Y \end{pmatrix} \right\|_1 \right] \\
 &\leq \hat{\mathbb{E}}_n \left[ \left\| \begin{pmatrix} 1+L \\ L \end{pmatrix} \right\|_\infty \cdot \left\| \begin{pmatrix} \eta_{\text{KCMC}}^T \boldsymbol{\psi}(T, X) - \eta_{\text{KCMC}}^{(\delta)T} \boldsymbol{\psi}(T, X) \\ \left( \frac{\pi_\beta(T|X)}{p_{\text{obs}}(T|X)} \right) Y - \left( \frac{\pi_{\beta^{(\delta)}}(T|X)}{p_{\text{obs}}(T|X)} \right) Y \end{pmatrix} \right\|_1 \right] \\
 &\leq (L+1) \cdot \left( \hat{\mathbb{E}}_n \left| \eta_{\text{KCMC}}^T \boldsymbol{\psi}(T, X) - \eta_{\text{KCMC}}^{(\delta)T} \boldsymbol{\psi}(T, X) \right| + \hat{\mathbb{E}}_n \left| \left( \frac{\pi_\beta(T|X)}{p_{\text{obs}}(T|X)} \right) Y - \left( \frac{\pi_{\beta^{(\delta)}}(T|X)}{p_{\text{obs}}(T|X)} \right) Y \right| \right) \\
 &\leq (L+1) \cdot \frac{2\delta}{3(L+1)} = \frac{2\delta}{3}, \tag{32}
 \end{aligned}$$

where  $J_{\eta_f}(h, r, t, x) := -h + \eta_f f_{t,x}^* \left( \frac{h-r}{\eta_f} \right)$ . The subgradients of  $J_{\eta_f}$  can be calculated as  $\partial_h J_{\eta_f}(h, r, t, x) := -1 + \partial f_{t,x}^* \left( \frac{h-r}{\eta_f} \right)$  and  $\partial_r J_{\eta_f}(h, r, t, x) := -\partial f_{t,x}^* \left( \frac{h-r}{\eta_f} \right)$  whose absolute value can be bounded by  $|1+L|$  and  $|L|$  respectively. Finally, by combining (31) and (32), we get

$$\hat{\mathbb{E}}_n \left| \ell_{\eta_f, \eta_{\text{KCMC}}, \beta}(Z) - \ell_{\eta_f^{(\delta)}, \eta_{\text{KCMC}}^{(\delta)}, \beta^{(\delta)}}(Z) \right| \leq \delta$$

with high probability as  $n \rightarrow \infty$ . Thus,  $\Theta^{(\delta)}$  is a  $\delta$ -covering of  $\Theta_\varepsilon$  satisfying  $\frac{\log |\Theta^{(\delta)}|}{n} \xrightarrow{P} 0$ . Thus, we conclude that

$$\sup_{\theta \in \Theta_\varepsilon} \left| \hat{\mathbb{E}}_n \ell_\theta(Z) - \mathbb{E} \ell_\theta(Z) \right| \xrightarrow{P} 0$$

by Lemma 16. ■

Finally, with the lemmas above, we can prove Theorem 6.

**Theorem 5 (Consistency of VC policy learning (full version))** *Assume policy class  $\{\pi_\beta(t|x) : \beta \in \mathcal{B}\}$  is VC and that Condition 2 is met. Then, we have  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) \xrightarrow{P} V_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*})$ .*

**Proof** Here, our aim is to prove the uniform convergence of the KCMC estimator over the policy class, i.e.

$$\sup_{\beta \in \mathcal{B}} \left| \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_\beta) - V_{\text{inf}}^{\text{KCMC}}(\pi_\beta) \right| \xrightarrow{P} 0.$$

Let  $\hat{\eta}(\beta) := \arg \min_{\eta \in H} \hat{\mathbb{E}}_n [\ell_{\beta, \eta}(Z)]$ . Let us also use Lemma 18 to take  $\varepsilon > 0$  so that

$$\sup_{\theta \in \Theta_\varepsilon} \left| \hat{\mathbb{E}}_n \ell_\theta(Z) - \mathbb{E} \ell_\theta(Z) \right| \xrightarrow{P} 0.$$

For this choice of  $\varepsilon$  and any  $\beta \in \mathcal{B}$ , let us define

$$\tilde{\eta}(\beta) := \frac{\varepsilon}{\varepsilon + \|\hat{\eta}(\beta) - \eta^*(\beta)\|} \cdot \hat{\eta}(\beta) + \frac{\|\hat{\eta}(\beta) - \eta^*(\beta)\|}{\varepsilon + \|\hat{\eta}(\beta) - \eta^*(\beta)\|} \cdot \eta^*(\beta) \tag{33}$$

so that it becomes a linear interpolation between  $\hat{\eta}(\beta)$  and  $\eta^*(\beta)$  that is always included in the  $\varepsilon$  neighborhood of  $\eta^*(\beta)$  so that  $\|\tilde{\eta}(\beta) - \eta^*(\beta)\| \leq \varepsilon$ . Since  $\eta \mapsto \ell_{\eta,\beta}(z)$  is convex for any  $z \in \mathcal{Z}$  and any  $\beta \in \mathcal{B}$ , we have

$$\hat{\mathbb{E}}_n[\ell_{\tilde{\eta}(\beta),\beta}(Z)] \leq \frac{\varepsilon}{\varepsilon + \|\hat{\eta}(\beta) - \eta^*(\beta)\|} \cdot \hat{\mathbb{E}}_n[\ell_{\hat{\eta}(\beta),\beta}(Z)] + \frac{\|\hat{\eta}(\beta) - \eta^*(\beta)\|}{\varepsilon + \|\hat{\eta}(\beta) - \eta^*(\beta)\|} \cdot \hat{\mathbb{E}}_n[\ell_{\eta^*(\beta),\beta}(Z)].$$

As the optimality of  $\hat{\eta}(\beta)$  with respect to  $\beta \mapsto \hat{\mathbb{E}}_n \ell_{\eta,\beta}(Z)$  implies  $0 \leq \hat{\mathbb{E}}_n[\ell_{\tilde{\eta}(\beta),\beta}(Z)] - \hat{\mathbb{E}}_n[\ell_{\hat{\eta}(\beta),\beta}(Z)]$ , we get

$$0 \leq \hat{\mathbb{E}}_n[\ell_{\eta^*(\beta),\beta}(Z)] - \hat{\mathbb{E}}_n[\ell_{\tilde{\eta}(\beta),\beta}(Z)].$$

Since  $(\eta^*(\beta), \beta)$  and  $(\tilde{\eta}(\beta), \beta)$  are included in  $\Theta_\varepsilon$ , we can use Lemma 18 to see

$$\begin{aligned} 0 &\leq \sup_{\beta \in \mathcal{B}} [\mathbb{E}[\ell_{\tilde{\eta}(\beta),\beta}(Z)] - \mathbb{E}[\ell_{\eta^*(\beta),\beta}(Z)]] \\ &\leq \sup_{\beta \in \mathcal{B}} \left[ \left( \mathbb{E}[\ell_{\tilde{\eta}(\beta),\beta}(Z)] - \hat{\mathbb{E}}_n[\ell_{\tilde{\eta}(\beta),\beta}(Z)] \right) - \left( \mathbb{E}[\ell_{\eta^*(\beta),\beta}(Z)] - \hat{\mathbb{E}}_n[\ell_{\eta^*(\beta),\beta}(Z)] \right) \right. \\ &\quad \left. - \left( \hat{\mathbb{E}}_n[\ell_{\eta^*(\beta),\beta}(Z)] - \hat{\mathbb{E}}_n[\ell_{\tilde{\eta}(\beta),\beta}(Z)] \right) \right] \\ &\leq \sup_{\beta \in \mathcal{B}} \left[ \left( \mathbb{E}[\ell_{\tilde{\eta}(\beta),\beta}(Z)] - \hat{\mathbb{E}}_n[\ell_{\tilde{\eta}(\beta),\beta}(Z)] \right) - \left( \mathbb{E}[\ell_{\eta^*(\beta),\beta}(Z)] - \hat{\mathbb{E}}_n[\ell_{\eta^*(\beta),\beta}(Z)] \right) \right] \\ &\leq 2 \sup_{(\eta,\beta) \in \Theta_\varepsilon} \left| \mathbb{E}[\ell_{\eta,\beta}(Z)] - \hat{\mathbb{E}}_n[\ell_{\eta,\beta}(Z)] \right| \\ &\stackrel{p}{\rightarrow} 0, \end{aligned}$$

which implies

$$\sup_{\beta \in \mathcal{B}} \left| \mathbb{E}[\ell_{\tilde{\eta}(\beta),\beta}(Z)] - \mathbb{E}[\ell_{\eta^*(\beta),\beta}(Z)] \right| \stackrel{p}{\rightarrow} 0.$$

Due to the uniform well-separatedness assumption in Condition 2, this leads to

$$\sup_{\beta \in \mathcal{B}} \|\tilde{\eta}(\beta) - \eta^*(\beta)\| \stackrel{p}{\rightarrow} 0,$$

and by definition of  $\tilde{\eta}$  (33), we have uniform consistency of  $\hat{\eta}(\beta)$ ,

$$\sup_{\beta \in \mathcal{B}} \|\hat{\eta}(\beta) - \eta^*(\beta)\| \stackrel{p}{\rightarrow} 0.$$

As  $\eta \mapsto \mathbb{E} \ell_{\eta,\beta}(z)$  has a Lipschitz constant  $\mathbb{E} \sqrt{\left| G_\varepsilon^{\partial_{\eta_f} \ell}(Z) \right|^2 + \left| G_\varepsilon^{\partial_{\eta_{\text{KCMC}}} \ell}(Z) \right|^2}$  that does not depend on  $\beta$  due to Lemma 17, we obtain

$$\sup_{\beta \in \mathcal{B}} \left| \hat{\mathbb{E}}_n [\ell_{\beta,\hat{\eta}(\beta)}(Z)] - \mathbb{E} [\ell_{\beta,\eta^*(\beta)}(Z)] \right| \stackrel{p}{\rightarrow} 0$$

i.e.

$$\sup_{\beta \in \mathcal{B}} \left| \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_\beta) - V_{\text{inf}}^{\text{KCMC}}(\pi_\beta) \right| \stackrel{p}{\rightarrow} 0.$$

Then, we have

$$\begin{aligned}
 0 &\geq V_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) - V_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*}) \\
 &\geq \left( V_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) - \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) \right) - \left( V_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*}) - \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*}) \right) \\
 &\quad + \left( \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) - \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*}) \right) \\
 &\geq \left( V_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) - \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) \right) - \left( V_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*}) - \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*}) \right) \\
 &\geq -2 \sup_{\beta \in \mathcal{B}} \left| \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\beta}) - V_{\text{inf}}^{\text{KCMC}}(\pi_{\beta}) \right| \xrightarrow{p} 0.
 \end{aligned}$$

Thus,

$$V_{\text{inf}}(\pi_{\hat{\beta}}) - V_{\text{inf}}(\pi_{\beta^*}) \xrightarrow{p} 0.$$

and

$$\left| \hat{V}_{\text{inf}}(\pi_{\hat{\beta}}) - V_{\text{inf}}(\pi_{\beta^*}) \right| \leq \left| \hat{V}_{\text{inf}}(\pi_{\hat{\beta}}) - V_{\text{inf}}(\pi_{\hat{\beta}}) \right| + \left| V_{\text{inf}}(\pi_{\hat{\beta}}) - V_{\text{inf}}(\pi_{\beta^*}) \right| \xrightarrow{p} 0,$$

which concludes the proof.  $\blacksquare$

### C.10 Proof of Theorem 7

To derive the asymptotic distribution of  $\hat{\theta} = (\hat{\eta}_f, \hat{\eta}_{\text{KCMC}})$ , we need the following regularity conditions in addition to Condition 1.

#### Condition 3 (Regularity of loss function III)

1.  $\mathbb{E}|\ell_{\theta^*}(Z)|^2 < \infty$ .
2.  $\theta \mapsto \mathbb{E}\ell_{\theta}(Z)$  is twice differentiable with positive definite Hessian at  $\theta = \theta^*$  so that

$$V := \nabla_{\theta} \nabla_{\theta}^T \mathbb{E}\ell_{\theta^*}(Z) \succ 0.$$

3.  $\theta \mapsto \ell_{\theta}(Z)$  is differentiable at  $\theta = \theta^*$  almost everywhere, i.e.,

$$\mathbb{E}\mathbf{1}_{\{z \in \mathcal{Z}: \theta \mapsto \ell_{\theta}(z) \text{ is not differentiable at } \theta = \theta^*\}}(Z) = 0.$$

4.  $f_{t,x}^*$  is uniformly Lipschitz continuous, i.e., there exists  $L > 0$  such that

$$\sup_{x \in \mathcal{X}, t \in \mathcal{T}, v \in \mathbb{R}} \sup_{\delta f^* \in \partial f_{t,x}^*} |\delta f^*(v)| \leq L.$$

5.  $\mathbb{E}|G^{\psi}(Z)|^2 < \infty$  and  $\mathbb{E}\left|G_{\varepsilon}^{f^*}(Z)\right|^2 < \infty$  for

$$G^{\psi}(z) := \|\psi(t, x)\|,$$

$$G_{\varepsilon}^{f^*}(z) := \sup_{\theta: \|\theta - \theta^*\| \leq \varepsilon} \left| f_{t,x}^* \left( \frac{\eta_{\text{KCMC}}^T \psi(t, x) - \left( \frac{\pi_{\beta}(t|x)}{p_{\text{obs}}(t|x)} \right) y}{\eta_f} \right) \right|$$

for some  $\varepsilon > 0$ .

6.  $\mathbb{E}|R(Z)|^2 < \infty$  for  $R(z) := \sup_{\beta \in \mathcal{B}} \left| \left( \frac{\pi_\beta(t|x)}{p_{\text{obs}}(t|x)} \right) y \right|$ .

With regularity conditions 3, we can apply the following lemma to immediately obtain the asymptotic normality of our estimator.

**Lemma 20** (*Vaart and Wellner, 1996, Example 3.2.12, 3.2.22*) *Assume  $\ell_\theta$  is Lipschitz in the neighborhood of  $\theta = \theta^*$  and is differentiable in quadratic mean at  $\theta = \theta^*$  with respect to parameter  $\theta$ , i.e. there exist  $\varepsilon > 0$ ,  $M : \mathcal{Z} \rightarrow \mathbb{R}$ , and  $\delta\ell_{\theta^*} : \mathcal{Z} \rightarrow \mathbb{R}^{D+1}$  such that*

$$|\ell_{\theta_1}(z) - \ell_{\theta_2}(z)| \leq M(z)\|\theta_1 - \theta_2\|$$

for any  $\theta_1, \theta_2 \in \{\theta : \|\theta - \theta^*\| \leq \varepsilon\}$ ,  $\mathbb{E}M(Z)^2 < \infty$ , and

$$\lim_{\theta \rightarrow \theta^*} \mathbb{E} \left| \frac{\ell_\theta(z) - \ell_{\theta^*}(z) - (\theta - \theta^*)^T \delta\ell_{\theta^*}(z)}{\|\theta - \theta^*\|} \right|^2 = 0. \quad (34)$$

Further assume that  $\theta \mapsto \mathbb{E}\ell_\theta(Z)$  is twice differentiable with positive definite Hessian at  $\theta = \theta^*$  so that  $V := \nabla_\theta \nabla_\theta^T \mathbb{E}\ell_{\theta^*}(Z) \succ 0$ . Then, for  $\hat{\theta} \xrightarrow{P} \theta^*$ , we have

$$\sqrt{n} \left[ \hat{\mathbb{E}}_n \ell_{\hat{\theta}}(Z) - \hat{\mathbb{E}}_n \ell_{\theta^*}(Z) \right] \xrightarrow{P} 0,$$

and

$$\sqrt{n} \left[ V(\hat{\theta} - \theta^*) + \hat{\mathbb{E}}_n \delta\ell_{\theta^*}(Z) \right] \xrightarrow{P} 0.$$

With this lemma, we prove Theorem 7 as follows:

**Theorem 6 (Asymptotic normality (full version))** *If loss function  $\ell_\theta(z)$  satisfies Condition 1 and Condition 3, then we have*

$$\sqrt{n} \left[ \hat{\mathbb{E}}_n \ell_{\hat{\theta}}(Z) - \hat{\mathbb{E}}_n \ell_{\theta^*}(Z) \right] \xrightarrow{P} 0,$$

and

$$\sqrt{n} \left[ V(\hat{\theta} - \theta^*) + \hat{\mathbb{E}}_n \delta\ell_{\theta^*}(Z) \right] \xrightarrow{P} 0.$$

therefore,

$$\sqrt{n}[\hat{\theta} - \theta^*] \xrightarrow{P} V^{-1} \sqrt{n} \hat{\mathbb{E}}_n [\delta\ell_{\theta^*}(Z)] \xrightarrow{d} N(0, V^{-1} J V^{-1}),$$

where  $V$  is defined as in Condition 3,  $J := \mathbb{E} [\delta\ell_{\theta^*}(Z) \delta\ell_{\theta^*}(Z)^T]$ , and  $\xrightarrow{d}$  implies the convergence in distribution.

**Proof** Here, we just have to show that the assumptions in Lemma 20 are satisfied. Due to Theorem 4, we know that  $\hat{\theta} \xrightarrow{P} \theta^*$ . We can take  $\delta\ell_{\theta^*} \in \partial\ell_{\theta^*}$  such that  $\mathbb{E}[\delta\ell_{\theta^*}] = 0$  and it satisfies the differentiability in quadratic mean (34). Now, we can take  $M(z) = G_\varepsilon^{\partial\ell} := \sup_{\theta: \|\theta - \theta^*\| \leq \varepsilon} \sup_{\delta\ell \in \partial\ell_\theta} \|\delta\ell(z)\|$ . Thus, it remains to show that

$$\mathbb{E} \left| G_\varepsilon^{\partial\ell}(Z) \right|^2 < \infty.$$

Indeed, using the same argument as Lemma 17, we have

$$\left| G_\varepsilon^{\partial_\theta \ell}(z) \right|^2 \leq \left| G_\varepsilon^{\partial_{\eta_f} \ell}(z) \right|^2 + \left| G_\varepsilon^{\partial_{\eta_{\text{KCMC}}} \ell}(z) \right|^2$$

for

$$G_\varepsilon^{\partial_{\eta_f} \ell}(z) := \sup_{\substack{\delta_{\eta_f} \ell \in \partial_{\eta_f} \ell_\theta(z) \\ \theta: \|\theta - \theta^*\| \leq \varepsilon}} |\delta_{\eta_f} \ell| \leq \gamma + G_\varepsilon^{f^*}(z) + L \left( \frac{(\theta_{\text{KCMC}}^* + \varepsilon) G^\psi(z) + R(z)}{\eta_f^* - \varepsilon} \right)$$

and

$$G_\varepsilon^{\partial_{\eta_{\text{KCMC}}} \ell}(z) := \sup_{\substack{\delta \ell \in \partial_{\eta_{\text{KCMC}}} \ell_\theta(z) \\ \theta: \|\theta - \theta^*\| \leq \varepsilon}} \|\delta \ell\| \leq (L + 1) G^\psi(z).$$

Since we can take  $\varepsilon$  satisfying  $\eta_f^* > \varepsilon > 0$  such that

$$\mathbb{E} \left| G_\varepsilon^{\partial_{\eta_f} \ell}(Z) \right|^2 \leq 4 \left\{ \gamma^2 + \mathbb{E} \left| G_\varepsilon^{f^*}(Z) \right|^2 + \frac{L^2 (\theta_{\text{KCMC}}^* + \varepsilon)^2 \mathbb{E} |G^\psi(Z)|^2}{|\eta_f^* - \varepsilon|^2} + \frac{L^2 \mathbb{E} |R(Z)|^2}{|\eta_f^* - \varepsilon|^2} \right\} < \infty$$

and

$$\mathbb{E} \left| G_\varepsilon^{\partial_{\eta_{\text{KCMC}}} \ell}(Z) \right|^2 \leq (L + 1)^2 \mathbb{E} \left| G^\psi(Z) \right|^2$$

due to Condition 3, we have  $\mathbb{E} \left| G_\varepsilon^{\partial_\theta \ell}(Z) \right|^2 < \infty$ . Therefore, we can apply Lemma 20 to our problem to conclude the proof.  $\blacksquare$

### C.11 Details of Example 8

**Proof** Due to the second order expansion in Condition 3, we have

$$\begin{aligned} & n \hat{\mathbb{E}}_n [\ell_{\hat{\theta}}(Z) - \ell_{\theta^*}(Z)] \\ &= \sqrt{n} \hat{\mathbb{E}}_n [\delta \ell_{\theta^*}(Z)]^T \sqrt{n} (\hat{\theta} - \theta^*) + \int_0^1 \sqrt{n} \hat{\mathbb{E}}_n [\delta \ell_{t(\hat{\theta} - \theta^*) + \theta^*}(Z) - \delta \ell_{\theta^*}(Z)]^T \sqrt{n} (\hat{\theta} - \theta^*) dt \\ &= \sqrt{n} \hat{\mathbb{E}}_n [\delta \ell_{\theta^*}(Z)]^T \sqrt{n} (\hat{\theta} - \theta^*) \\ &\quad + \int_0^1 \sqrt{n} (\hat{\mathbb{E}}_n - \mathbb{E}) \left[ \delta \ell_{t(\hat{\theta} - \theta^*) + \theta^*}(Z) - \delta \ell_{\theta^*}(Z) \right]^T \sqrt{n} (\hat{\theta} - \theta^*) dt \\ &\quad + \int_0^1 \sqrt{n} \mathbb{E} \left[ \delta \ell_{t(\hat{\theta} - \theta^*) + \theta^*}(Z) - \delta \ell_{\theta^*}(Z) \right]^T \sqrt{n} (\hat{\theta} - \theta^*) dt \\ &= \sqrt{n} \hat{\mathbb{E}}_n [\delta \ell_{\theta^*}(Z)]^T \sqrt{n} (\hat{\theta} - \theta^*) \\ &\quad + \int_0^1 o_p(1) \sqrt{n} (\hat{\theta} - \theta^*) dt \\ &\quad + \int_0^1 \sqrt{n} \mathbb{E} \left[ \delta \ell_{t(\hat{\theta} - \theta^*) + \theta^*}(Z) - \delta \ell_{\theta^*}(Z) \right]^T \sqrt{n} (\hat{\theta} - \theta^*) dt \\ &= \sqrt{n} \hat{\mathbb{E}}_n [\delta \ell_{\theta^*}(Z)]^T \sqrt{n} (\hat{\theta} - \theta^*) + o_p(1) \cdot O_p(\sqrt{n} \|\hat{\theta} - \theta^*\|) \end{aligned}$$

$$\begin{aligned}
 & + \int_0^1 \sqrt{n} \left\{ \nabla_{\theta} \mathbb{E}[\ell_{t(\hat{\theta}-\theta^*)+\theta^*}(Z)]^T - \nabla_{\theta} \mathbb{E}[\ell_{\theta^*}(Z)]^T \right\} \sqrt{n}(\hat{\theta} - \theta^*) dt \\
 & = \sqrt{n} \hat{\mathbb{E}}_n[\delta\ell_{\theta^*}(Z)]^T \sqrt{n}(\hat{\theta} - \theta^*) + o_p(1) + \int_0^1 \sqrt{n} \left( (\hat{\theta} - \theta^*)^T tV + o(\|\hat{\theta} - \theta^*\|) \right) \sqrt{n}(\hat{\theta} - \theta^*) dt \\
 & \xrightarrow{p} -\sqrt{n} (\hat{\theta} - \theta^*)^T V \sqrt{n}(\hat{\theta} - \theta^*) + \frac{1}{2} \sqrt{n} (\hat{\theta} - \theta^*)^T V \sqrt{n}(\hat{\theta} - \theta^*) \\
 & = -\frac{1}{2} \sqrt{n} (\hat{\theta} - \theta^*)^T V \sqrt{n}(\hat{\theta} - \theta^*),
 \end{aligned}$$

where we made an approximation  $\sqrt{n} (\hat{\mathbb{E}}_n - \mathbb{E}) \left[ \delta\ell_{t(\hat{\theta}-\theta^*)+\theta^*}(Z) - \delta\ell_{\theta^*}(Z) \right] \xrightarrow{p} 0$ . <sup>24</sup>  $\blacksquare$

## Appendix D. Derivation of Hessian for the box constraints

Here, we derive the Hessian of  $f^*$  in the case of box constraints. For conjugate function

$$f_{t,x}^*(v) = \begin{cases} a(t,x)v & \text{if } v < 0, \\ 0 & \text{if } v = 0, \\ b(t,x)v & \text{if } v > 0, \end{cases}$$

and its subgradient

$$\partial f_{t,x}^*(v) = \begin{cases} a(t,x) & \text{if } v < 0, \\ [a(t,x), b(t,x)] & \text{if } v = 0, \\ b(t,x) & \text{if } v > 0, \end{cases}$$

the gradient of the dual objective <sup>25</sup> is

$$\begin{aligned}
 & \nabla_{\eta_{\text{KCMC}}} \left\{ \eta_{\text{KCMC}}^T \mathbb{E}[\boldsymbol{\psi}] - \mathbb{E} \left[ f_{T,X}^* (\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r) \right] \right\} \\
 & = \mathbb{E} \left[ (1 - \partial f_{T,X}^* (\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r)) \boldsymbol{\psi} \right].
 \end{aligned}$$

Therefore, the second-order derivatives are:

$$\nabla_{\eta_{\text{KCMC}}} \nabla_{\eta_{\text{KCMC}}}^T \left\{ \eta_{\text{KCMC}}^T \mathbb{E}[\boldsymbol{\psi}] - \eta_f \mathbb{E} \left[ f_{T,X}^* (\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r) \right] \right\}$$

24. Though this approximation is not justified by our previous asymptotic theorem, it is valid when  $\{\delta\ell_{\theta} : \|\theta - \theta^*\| < \varepsilon\}$  is a Donsker class, for which the uniform central limit theorem holds. Though proving the Donsker property for general  $f_{x,t}$  is difficult, we can still show that Example 5 with  $f_{t,x}^0 = f^0$  not depending on  $t$  and  $x$  satisfies this condition. To sketch the proof, we first show that  $\partial\ell_{\theta}(z)$  consists of terms that are VC classes by using reparametrization  $\nu = (1/\eta_f, \eta_{\text{KCMC}}/\eta_f)$  and stability of VC classes (Vaart and Wellner, 1996, 2.6.18) along with the fact that  $f_{t,x}^*(v) = \max\{f^{0*}(v), \min\{a_{t,x}v - f^0(a_{t,x}), b_{t,x}v - f^0(b_{t,x})\}\}$ ,  $\partial f_{t,x}^*(v) = \min\{b_{t,x}, \max\{a_{t,x}, \partial f^{0*}(v)\}\}$  and that  $\partial f^*$  is monotone. Then, we use the fact that VC classes have bounded uniform entropy and are Donsker class (Vaart and Wellner, 1996, 2.5.2, 2.6.7), and their permanence (Vaart and Wellner, 1996, 2.10.6, 2.10.20, 2.10.23) to show  $\{\delta\ell_{\theta} : \|\theta - \theta^*\| < \varepsilon\}$  is a Donsker class.
25. For the box constraints, it can be shown that the dual objective becomes  $-\gamma\eta_f + \eta_{\text{KCMC}}^T \mathbb{E}[\boldsymbol{\psi}] - \mathbb{E} \left[ f_{T,X}^* (\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r) \right]$  and that its supremum is reached when  $\eta_f \rightarrow 0$ . Furthermore, when we extend the domain of the dual objective from  $\eta_f > 0$  to  $\eta_f \geq 0$ , we know that  $\lim_{n \rightarrow \infty} \mathbb{P}(\hat{\eta}_f = 0) \rightarrow 1$ , so  $\hat{\eta}_f$  does not need any asymptotic analysis.

$$\begin{aligned}
 &= \nabla_{\eta_{\text{KCMC}}} \mathbb{E} \left[ (1 - \partial f_{T,X}^* (\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r)) \boldsymbol{\psi}^T \right] \\
 &= -\mathbb{E} \left[ \left\{ \nabla_{\eta_{\text{KCMC}}} \mathbb{E} \left[ \partial f_{T,X}^* (\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r) \mid T, X \right] \right\} \boldsymbol{\psi}^T \right] \\
 &= -\mathbb{E} \left[ \left\{ \nabla_{\eta_{\text{KCMC}}} \mathbb{E} \left[ a(T, X) \mathbf{1}_{\mathbb{R}_-} (\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r) + b(T, X) \mathbf{1}_{\mathbb{R}_+} (\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r) \mid T, X \right] \right\} \boldsymbol{\psi}^T \right] \\
 &= -\mathbb{E} \left[ \left\{ \nabla_{\eta_{\text{KCMC}}} a(T, X) \mathbb{P} \left[ (\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r) < 0 \mid T, X \right] \right. \right. \\
 &\quad \left. \left. + \nabla_{\eta_{\text{KCMC}}} b(T, X) \mathbb{P} \left[ (\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r) > 0 \mid T, X \right] \right\} \boldsymbol{\psi}^T \right] \\
 &= -\mathbb{E} \left[ \left\{ \nabla_{\eta_{\text{KCMC}}} a(T, X) \mathbb{P} \left[ \eta_{\text{KCMC}}^T \boldsymbol{\psi} < r \mid T, X \right] + \nabla_{\eta_{\text{KCMC}}} b(T, X) \mathbb{P} \left[ \eta_{\text{KCMC}}^T \boldsymbol{\psi} > r \mid T, X \right] \right\} \boldsymbol{\psi}^T \right] \\
 &= -\mathbb{E} \left[ \boldsymbol{\psi} \boldsymbol{\psi}^T \{ b(T, X) - a(T, X) \} p_r(\eta_{\text{KCMC}}^T \boldsymbol{\psi} \mid T, X) \right],
 \end{aligned}$$

where  $p_r(\cdot \mid t, x)$  is the probability density function of  $r(Y, T, X)$  conditioned on  $T = t, X = x$ .

## References

- MOSEK ApS. Mosek optimization suite, 2019.
- Pierre-Cyril Aubin-Frankowski and Zoltán Szabó. Hard shape-constrained kernel machines. Advances in Neural Information Processing Systems, 33:384–395, 2020.
- Andrew Bennett, Nathan Kallus, Miruna Oprescu, Wen Sun, and Kaiwen Wang. Efficient and sharp off-policy evaluation in robust markov decision processes. Advances in Neural Information Processing Systems, 37:112962–113000, 2024.
- Gilles Blanchard, Olivier Bousquet, and Laurent Zwald. Statistical properties of kernel principal component analysis. Machine Learning, 66(2):259–294, 2007.
- Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. Convex optimization. Cambridge university press, 2004.
- David Bruns-Smith and Angela Zhou. Robust fitted-q-evaluation and iteration under sequentially exogenous unobserved confounders. arXiv preprint arXiv:2302.00662, 2023.
- Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters, 2018.
- John M Danskin. The theory of max-min, with applications. SIAM Journal on Applied Mathematics, 14(4):641–664, 1966.
- Steven Diamond and Stephen Boyd. CVXPY: A Python-embedded modeling language for convex optimization. Journal of Machine Learning Research, 17(83):1–5, 2016.
- A. Domahidi, E. Chu, and S. Boyd. ECOS: An SOCP solver for embedded systems. In European Control Conference (ECC), pages 3071–3076, 2013.
- Jacob Dorn and Kevin Guo. Sharp sensitivity analysis for inverse propensity weighting via quantile balancing. Journal of the American Statistical Association, 2022.

- Jacob Dorn, Kevin Guo, and Nathan Kallus. Doubly-valid/doubly-sharp sensitivity analysis for causal inference with unmeasured confounding. arXiv preprint arXiv:2112.11449, 2021.
- Arthur Gretton, Olivier Bousquet, Alex Smola, and Bernhard Schölkopf. Measuring statistical dependence with hilbert-schmidt norms. In International conference on algorithmic learning theory, pages 63–77. Springer, 2005.
- Jason Hartford, Greg Lewis, Kevin Leyton-Brown, and Matt Taddy. Deep iv: A flexible approach for counterfactual prediction. In International Conference on Machine Learning, pages 1414–1423. PMLR, 2017.
- Keisuke Hirano and Guido W Imbens. Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. Health Services and Outcomes research methodology, 2(3):259–278, 2001.
- Keisuke Hirano, Guido W Imbens, and Geert Ridder. Efficient estimation of average treatment effects using the estimated propensity score. Econometrica, 71(4):1161–1189, 2003.
- Kei Ishikawa and Niao He. Kernel conditional moment constraints for confounding robust inference. In International Conference on Artificial Intelligence and Statistics, pages 650–674. PMLR, 2023.
- Ying Jin, Zhimei Ren, and Zhengyuan Zhou. Sensitivity analysis under the  $f$ -sensitivity models: Definition, estimation and inference. arXiv preprint arXiv:2203.04373, 2022.
- Nathan Kallus. Balanced policy evaluation and learning. Advances in neural information processing systems, 31, 2018.
- Nathan Kallus and Angela Zhou. Confounding-robust policy improvement. Advances in neural information processing systems, 31, 2018.
- Nathan Kallus and Angela Zhou. Confounding-robust policy evaluation in infinite-horizon reinforcement learning. Advances in Neural Information Processing Systems, 33:22293–22304, 2020.
- Nathan Kallus and Angela Zhou. Minimax-optimal policy learning under unobserved confounding. Management Science, 67(5):2870–2890, 2021.
- Nathan Kallus, Xiaojie Mao, and Masatoshi Uehara. Causal inference under unmeasured confounding with negative controls: A minimax learning approach. arXiv preprint arXiv:2103.14029, 2021.
- Daphne Koller and Nir Friedman. Probabilistic graphical models: principles and techniques. MIT press, 2009.
- Sadanori Konishi and Genshiro Kitagawa. Information criteria and statistical modeling. 2008.

- Heiner Kremer, Jia-Jie Zhu, Krikamol Muandet, and Bernhard Schölkopf. Functional generalized empirical likelihood estimation for conditional moment restrictions. In International Conference on Machine Learning, pages 11665–11682. PMLR, 2022.
- Weiwei Liu, S Janet Kuramoto, and Elizabeth A Stuart. An introduction to sensitivity analysis for unobserved confounding in nonexperimental prevention research. Prevention science, 14(6):570–580, 2013.
- 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.
- Afsaneh Mastouri, Yuchen Zhu, Limor Gultchin, Anna Korba, Ricardo Silva, Matt Kusner, Arthur Gretton, and Krikamol Muandet. Proximal causal learning with kernels: Two-stage estimation and moment restriction. In International Conference on Machine Learning, pages 7512–7523. PMLR, 2021.
- Charles A Micchelli, Yuesheng Xu, and Haizhang Zhang. Universal kernels. Journal of Machine Learning Research, 7(12), 2006.
- Krikamol Muandet, Wittawat Jitkrittum, and Jonas Kübler. Kernel conditional moment test via maximum moment restriction. In Conference on Uncertainty in Artificial Intelligence, pages 41–50. PMLR, 2020a.
- Krikamol Muandet, Arash Mehrjou, Si Kai Lee, and Anant Raj. Dual instrumental variable regression. Advances in Neural Information Processing Systems, 33:2710–2721, 2020b.
- Miruna Oprescu, Jacob Dorn, Marah Ghoummaid, Andrew Jesson, Nathan Kallus, and Uri Shalit. B-learner: Quasi-oracle bounds on heterogeneous causal effects under hidden confounding. In International Conference on Machine Learning, pages 26599–26618. PMLR, 2023.
- Judea Pearl. Causality. Cambridge university press, 2009.
- Adrián Pérez-Suay, Valero Laparra, Gonzalo Mateo-García, Jordi Muñoz-Marí, Luis Gómez-Chova, and Gustau Camps-Valls. Fair kernel learning. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pages 339–355. Springer, 2017.
- Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. Elements of causal inference: foundations and learning algorithms. The MIT Press, 2017.
- Paul R Rosenbaum. Overt bias in observational studies. In Observational studies, pages 71–104. Springer, 2002.
- Paul R Rosenbaum, PR Rosenbaum, and Briskman. Design of observational studies, volume 10. Springer, 2010.
- 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.
- Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Kernel principal component analysis. In International conference on artificial neural networks, pages 583–588. Springer, 1997.
- 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.
- Rahul Singh. Kernel methods for unobserved confounding: Negative controls, proxies, and instruments. arXiv preprint arXiv:2012.10315, 2020.
- Rahul Singh, Maneesh Sahani, and Arthur Gretton. Kernel instrumental variable regression. Advances in Neural Information Processing Systems, 32, 2019.
- Rahul Singh, Liyuan Xu, and Arthur Gretton. Kernel methods for causal functions: dose, heterogeneous and incremental response curves. Biometrika, 111(2):497–516, 2024.
- Matthew Staib and Stefanie Jegelka. Distributionally robust optimization and generalization in kernel methods. Advances in Neural Information Processing Systems, 32, 2019.
- Zhiqiang Tan. A distributional approach for causal inference using propensity scores. Journal of the American Statistical Association, 101(476):1619–1637, 2006.
- Aad W Vaart and Jon A Wellner. Weak convergence. In Weak convergence and empirical processes, pages 16–28. Springer, 1996.
- Sara Van de Geer. Empirical Processes in M-estimation, volume 6. Cambridge university press, 2000.
- Sara Van de Geer. Empirical processes theory. 2020. URL [www.stat.math.ethz.ch/~geer/empirical-process2020.pdf](http://www.stat.math.ethz.ch/~geer/empirical-process2020.pdf).
- 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.
- Kun Zhang, Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. Kernel-based conditional independence test and application in causal discovery. arXiv preprint arXiv:1202.3775, 2012.
- Rui Zhang, Masaaki Imaizumi, Bernhard Schölkopf, and Krikamol Muandet. Instrumental variable regression via kernel maximum moment loss. Journal of Causal Inference, 11(1):20220073, 2023.
- Qingyuan Zhao, Dylan S Small, and Bhaswar B Bhattacharya. Sensitivity analysis for inverse probability weighting estimators via the percentile bootstrap. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 81(4):735–761, 2019.

Jia-Jie Zhu, Wittawat Jitkrittum, Moritz Diehl, and Bernhard Schölkopf. Worst-case risk quantification under distributional ambiguity using kernel mean embedding in moment problem. In 2020 59th IEEE Conference on Decision and Control (CDC), pages 3457–3463. IEEE, 2020.