

---

# Kernel Conditional Moment Constraints for Confounding Robust Inference

---

**Kei Ishikawa**  
ETH Zürich  
kishikawa@student.ethz.ch

**Niao He**  
ETH Zürich  
niao.he@inf.ethz.ch

## 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 over 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. It can be shown that our estimator contains the recently proposed sharp estimator by Dorn and Guo (2022) as a special case, and our method enables a novel extension of the classical marginal sensitivity model using  $f$ -divergence. To construct our estimator, we leverage the kernel method to obtain a tractable approximation to the conditional moment constraints, which traditional non-sharp estimators failed to take into account. In the theoretical analysis, we provide a condition for the choice of the kernel which guarantees no specification error that biases the lower bound estimation. Furthermore, we provide consistency guarantee of policy evaluation and extend the result to policy learning. In the experiments with synthetic and real-world data, we demonstrate the effectiveness of the proposed method.

## 1 INTRODUCTION

The offline contextual bandit is a simple but powerful model for decision-making with a wide range of applications such as data-driven personalized medical treatment, recommendations, and advertisements on online platforms. In the

evaluation of its policy value, the inverse probability weighting (IPW) method (Hirano and Imbens, 2001; Hirano et al., 2003) or its variant is commonly used. This method relies on a so-called *unconfoundedness* assumption, which essentially requires full observability of all relevant variables so that there exist no unobserved variables that influence the selection of action and resulting reward (Rubin, 1974). However, in practice, such an assumption can easily be violated due to the existence of unobserved confounders that are not recorded in the logged data.

A common way to address this problem is resorting to the worst-case lower bound of the policy value, namely, we minimize the policy value over a plausible uncertainty set that contains all the possible confounding situations. With such a lower bound, we can make an informed decision that is robust to confounding. The estimation and inference of such a lower bound are called sensitivity analysis and it has been extensively studied over the years (Rosenbaum, 2002; Tan, 2006; Rosenbaum et al., 2010; Liu et al., 2013). Among a wide range of existing sensitivity models, a popular choice is the marginal sensitivity model by Tan (2006) and its extensions. Recently, Zhao et al. (2019) introduced an elegant algorithm for Tan’s marginal sensitivity model using the linear fractional programming, which has revitalized the study of this model. This approach was further extended to policy learning in Kallus and Zhou (2018, 2021).

However, these sensitivity analysis methods rely on algorithms using linear programming that finds an overly conservative lower bound of policy value. This is a fundamental problem, as these loose lower bound estimators are only guaranteed to be lower than or equal to the true lower bound of the uncertainty set, but they are not necessarily the consistent estimator of the true lower bound. Even so, these algorithms have been widely adopted for their tractability. To obtain a sharp lower bound, conditional moment constraints, which consist of infinite-dimensional linear constraints, must be leveraged. Recently, Dorn and Guo (2022) analyzed these constraints and characterized a sharp lower bound of Tan’s marginal sensitivity model using a conditional quantile function of the reward distribution. With this characterization, they proposed the first tractable algorithm to obtain the sharp lower bound that converges to the true

lower bound of the policy value.

In this paper, we address the same problem of sharp estimation from a new perspective. Instead of using the conditional quantile function, we employ the kernel method (Schölkopf et al., 2002), a rich and flexible modeling paradigm in machine learning. We develop a tractable kernel approximation of the conditional moment constraints and propose an efficient algorithm to obtain the sharp lower bound.

**Our contributions.** We summarize our contributions in several aspects below.

First, we extend the existing sensitivity analysis models by considering uncertainty sets characterized by more general convex constraints. Our model includes the original sensitivity model by Tan (2006) as the special case but it also includes a new f-sensitivity model that extends Tan (2006)’s sensitivity model using f-divergence.

Second, we provide efficient algorithms based on the kernel method and low-rank approximation to obtain sharp estimators of the worst-case lower bound for the extended model. Our new estimator is very general and it includes the previous sharp estimator by Dorn and Guo (2022) as a special case. Using the duality of the associated convex optimization problem, we further identify conditions for zero specification error guarantees and establish consistency guarantees of our estimator in policy evaluation.

Third, we show that our method can naturally be extended to policy learning, as it offers a very simple way to compute the policy gradient. This is an advantage of our estimator compared to the previous sharp estimator (Dorn and Guo, 2022), which does not offer the possibility of policy learning. We provide a consistency guarantee for policy learning with a sharp lower bound, which is similar to the guarantee for the non-sharp estimator by Kallus and Zhou (2018, 2021).

Last but not least, we demonstrate the effectiveness of imposing the kernel conditional moment constraints in several numerical experiments on both synthetic and real-world data. We cover a wide range of problems in sensitivity analysis such as the generalized sensitivity models defined with f-divergence and policy learning, and our estimator consistently outperforms the conventional non-sharp estimators in these settings.

**Related work.** Similar to our paper, Kremer et al. (2022) used the kernel method for parameter estimation of models characterized by conditional moment restrictions. They solved the dual of their original problem by using the dual representation of the  $L_2$ -norm of the conditional moment. Though we solve a primal problem in this paper, we take great advantage of such a dual formulation in our theoretical analysis. Muandet et al. (2020) considered hypothesis testing for conditional moment conditions. They constructed their test statistic using a quadratic form of ker-

nel matrix similar to the one we use. The idea of using the kernel method to impose constraints has also been explored in other contexts such as fair regression (Pérez-Suay et al., 2017), distributionally-robust optimization (Staub and Jegelka, 2019), worst-case risk quantification (Zhu et al., 2020), and shape constraints to derivatives (Aubin-Frankowski and Szabó, 2020). Recently, the kernel method has found various novel applications in causal inference, including instrumental variable regression (Singh et al., 2019), negative controls (Singh, 2020; Kallus et al., 2021; Mastouri et al., 2021), and conditional mean squared error minimization for policy evaluation (Kallus, 2018).

## 2 BACKGROUNDS AND PROBLEM SETTINGS

### 2.1 Confounded Offline Contextual Bandits

Confounded offline contextual bandits are an extension of the standard offline contextual bandits that have an additional unobserved confounding variable. We are interested in evaluating the value of policy  $\pi$  from the offline data following base policy  $\pi_{\text{base}}$ , which is generated according to the following model:

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

where only  $Y$ ,  $T$ , and  $X$  are observable and policy  $\pi_{\text{base}}(t|x, u)$  is unknown. Action  $t \in \mathcal{T}$  is chosen by the (stochastic) base policy given context  $X \in \mathcal{X}$  and unobserved variable  $U \in \mathcal{U}$ . Reward  $Y$  is randomly generated conditionally on the values of  $T$ ,  $X$ , and  $U$ .

In the offline evaluation of policy  $\pi$ , we are interested in the expectation of  $Y$  under the modified process of (1) where  $\pi_{\text{base}}(t|x, u)$  is replaced by  $\pi(t|x)$ . Thus, the desired policy value of  $\pi$  can be written as

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

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. For simplicity of notations, we denote  $\mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X, U)} [f(Y, T, X, U)]$  as the expectations of  $f(Y, T, X, U)$  under generative process (1) and  $\mathbb{E}_{T \sim \pi(\cdot|X)} [f(Y, T, X, U)]$  as its modification where  $\pi_{\text{base}}$  is replaced by  $\pi$ . Hereafter, we assume that  $\pi_{\text{base}}(T|X, U) > 0$  holds almost surely so that the inverse probability weights are always well-defined.

In unconfounded offline contextual bandits, we can use the

inverse probability weighting (IPW) estimator

$$\hat{V}_{\text{IPW}}(\pi) := \frac{1}{n} \sum_{i=1}^n \left( \frac{\pi(T_i|X_i)}{\hat{\pi}_{\text{base}}(T_i|X_i)} \right) Y_i$$

with estimated base policy  $\hat{\pi}_{\text{base}}$  to evaluate the policy  $\pi$  consistently. However, when  $\pi_{\text{base}}$  depends on  $U$ , we can no longer construct such a consistent estimator, as the observable variables are only  $Y$ ,  $T$ , and  $X$ , and any valid estimator must depend only on them.

To indicate a part of model (1) that can be approximated by the offline data, we use  $p_{\text{obs}}$  to indicate the observable distribution of (1) such that

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

Similarly,  $p_{\text{obs}}(t|x)$  and  $p_{\text{obs}}(x)$  denote the corresponding conditional and marginal distributions, and  $\mathbb{E}_{\text{obs}}[f(Y, T, X)]$  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, we use abbreviation  $\mathbb{E}_{\text{obs}}[f|t, x]$  to represent conditional expectation  $\mathbb{E}_{\text{obs}}[f(Y, T, X)|T = t, X = x]$ . Hereafter, these observable distributions are assumed to be available for constructing estimators.

## 2.2 Uncertainty Sets of Base Policies

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

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

In the following, we list a few types of constraints used for the construction of the uncertainty sets.

### Box Constraints

The box constraints have been widely adopted in the sensitivity analysis, and they can be written as

$$a_{\pi}(t, x) \leq \pi_{\text{base}}(t|x, u) \leq b_{\pi}(t, x) \quad (2)$$

for some  $a_{\pi}(t, x)$  and  $b_{\pi}(t, x)$ . This assumption is used in the well-known marginal sensitivity model by Tan (2006) as well as many of its extensions (Zhao et al., 2019; Kallus and Zhou, 2018; Dorn and Guo, 2022). Tan (2006) considered a binary action space and assumed that the odds ratio of

observational conditional probability  $p_{\text{obs}}(t|x)$  and the true base policy  $\pi_{\text{base}}(t|x, u)$  is not too far from 1 so that

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

As  $p_{\text{obs}}(t|x)$  can be estimated from the observational data, we can enforce such constraints by choosing  $a_{\pi}$  and  $b_{\pi}$  in (2) as  $a_{\pi}(t, x) = 1/(1 + \Gamma(1/p_{\text{obs}}(t|x) - 1))$  and  $b_{\pi}(t, x) = 1/(1 + \Gamma^{-1}(1/p_{\text{obs}}(t|x) - 1))$ .

### f-divergence Constraint

The f-divergence is a measure of dissimilarity between two distributions. For probability mass function (or density function)  $p(t)$  and  $q(t)$ , the f-divergence between them is defined as  $D_f[p||q] := \int f\left(\frac{p(t)}{q(t)}\right) q(t) dt$  for some convex function  $f: \mathbb{R} \rightarrow \mathbb{R}$  satisfying  $f(1) = 0$ .<sup>1</sup> It is a rich class of divergence between probability distributions that includes many divergences such as the Kullback–Leibler (KL) divergence. Using the f-divergence, we introduce a new class of sensitivity assumption<sup>2</sup>

$$\mathbb{E}_{X, U} [D_f[p_{\text{obs}}(t|X)||\pi_{\text{base}}(t|X, U)]] \leq \gamma, \quad (3)$$

where the expectation  $\mathbb{E}_{X, U}$  is taken with respect to  $p(x, u)$  in (1) regardless of the policy. By encoding the proximity of  $\pi_{\text{base}}(t|x, u)$  from  $p_{\text{obs}}(t|x)$  using the f-divergence instead of the box constraints, we can construct a flexible class of the uncertainty sets. As we see later, this formulation is computationally convenient, as it can be expressed as simple expectation

$$\begin{aligned} & \mathbb{E}_{X, U} [D_f[p_{\text{obs}}(t|X)||\pi_{\text{base}}(t|X, U)]] \\ &= \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X, U)} \left[ f\left( \frac{p_{\text{obs}}(T|X)}{\pi_{\text{base}}(T|X, U)} \right) \right]. \end{aligned}$$

Compared to the marginal sensitivity model (Tan, 2006) that imposes uniform bounds on the odds ratio  $p_{\text{obs}}(t|x)/\pi_{\text{base}}(t|x, u)$  for any  $x, u$ , our f-sensitivity model upper bounds its average deviations from 1 (i.e. the unconfounded case). Thus, when the odds ratio is locally very far from 1 around some  $x, u$  but is close to 1 elsewhere, the f-sensitivity model is a more reasonable choice than the conventional model.

### Conditional f-constraint

The above f-sensitivity model can be extended to an even more general case by letting convex function  $f$  depend on  $T$  and  $X$ . 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

<sup>1</sup>We use  $\int$  because our method can handle both discrete and continuous treatment spaces in the same way.

<sup>2</sup>Jin et al. (2022) proposed a similar uncertainty set based on f-divergence, but their model is different from ours. See more discussion in the supplementary material.

fixed  $t \in \mathcal{T}$  and  $x \in \mathcal{X}$ . Then, we introduce a conditional f-constraint defined as

$$\mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X,U)} \left[ f_{T,X} \left( \frac{p_{\text{obs}}(T|X)}{\pi_{\text{base}}(T|X,U)} \right) \right] \leq \gamma. \quad (4)$$

Clearly, this uncertainty set generalizes f-divergence constraint (3). Moreover, this model contains the box constraint (2) as a special case. By choosing

$$f_{t,x}(\tilde{w}) = \begin{cases} 0 & \text{if } a_{\tilde{w}}(t,x) \leq \tilde{w} \leq b_{\tilde{w}}(t,x) \\ \infty & \text{otherwise} \end{cases} \quad (5)$$

for  $a_{\tilde{w}}(t,x) = p_{\text{obs}}(t|x)/b_{\pi}(t,x)$  and  $b_{\tilde{w}}(t,x) = p_{\text{obs}}(t|x)/a_{\pi}(t,x)$ , it becomes equivalent to box constraints (5). To provide a systematic treatment of different types of uncertainty sets and unify the theoretical analysis, we will hereafter assume that the uncertainty sets of inverse probability weights always have some conditional f-constraint unless otherwise specified.

### 2.3 Relaxed Uncertainty Sets of Inverse Probability Weights

Let us introduce re-parametrization  $w(y,t,x) = \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X,U)} \left[ \frac{1}{\pi_{\text{base}}(T|X,U)} |y,t,x \right]$  to obtain tractable uncertainty sets. The uncertainty set for  $\pi_{\text{base}}$  described earlier requires reparametrized weight  $w(y,t,x)$  to satisfy the following two conditions:

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

and

$$w(y,t,x) = \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X,U)} \left[ \frac{1}{\pi_{\text{base}}(T|X,U)} |y,t,x \right]$$

for some proper policy  $\pi_{\text{base}}$ . In general, both conditions are intractable. Therefore, we will consider the relaxation of these conditions.

#### Relaxation of the Conditional f-constraint

Let us first consider conditional f-constraint (4). With Jensen's inequality, we have

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

Therefore, we can relax the condition (4) to

$$\mathbb{E}_{\text{obs}} [f_{T,X} (p_{\text{obs}}(T|X)w(Y,T,X))] \leq \gamma. \quad (6)$$

#### Relaxation of the Distributional Constraints

Now we consider the relaxation of the second constraint, i.e., there exists proper underlying distribution  $\pi_{\text{base}}(t|x,u)$  that yields  $w(y,t,x)$ . This constraint is usually relaxed to a one-dimensional linear constraint in previous work. Here, we present a tighter relaxation using infinite-dimensional linear constraints called *conditional moment constraints*. In the following, we present these two types of relaxation.

First, we describe the simple relaxation adopted in previous work (Zhao et al., 2019; Kallus and Zhou, 2018, 2021). When action space  $\mathcal{T}$  is discrete and finite, the distributional constraint can be relaxed to

$$\mathbb{E}_{\text{obs}} [\mathbb{1}_{T=t} w(Y,T,X)] = 1 \text{ for any } t \in \mathcal{T} \quad (7)$$

and

$$w(y,t,x) \geq 0 \text{ for any } y \in \mathcal{Y}, t \in \mathcal{T}, \text{ and } x \in \mathcal{X} \quad (8)$$

where  $\mathbb{1}$  denotes the indicator function for event  $A$ . Following the naming convention in Dorn and Guo (2022), we will call constraint (7) the ZSB constraint after the authors of Zhao et al. (2019).

Combining the above with the relaxation of the conditional f-constraint as in (6), the following uncertainty set with the ZSB constraint can be defined:

$$\mathcal{W}_{f_{t,x}}^{\text{ZSB}} := \left\{ w \geq 0 : \begin{array}{l} \mathbb{E}_{X,U} [f_{T,X} (p_{\text{obs}}(T|X)w)] \leq \gamma, \\ \mathbb{E}_{\text{obs}} [\mathbb{1}_{T=t} w] = 1 \text{ for any } t \in \mathcal{T} \end{array} \right\}.$$

For this uncertainty set, the associated lower bound is

$$V_{\text{inf}}^{\text{ZSB}}(\pi) := \inf_{w \in \mathcal{W}_{f_{t,x}}^{\text{ZSB}}} \mathbb{E}_{\text{obs}} [w(Y,T,X)\pi(T|X)Y]. \quad (9)$$

Now, we discuss the other relaxation based on conditional moment constraints. It can be shown that it is possible to relax the distributional constraints to

$$\mathbb{E}_{\text{obs}} [w(Y,T,X)|T=t,X=x] \cdot p_{\text{obs}}(t|x) = 1 \quad (10)$$

for any  $t \in \mathcal{T}$  and  $x \in \mathcal{X}$  plus a non-negativity constraint (8). The derivation of the above constraints is deferred to the supplementary material. Note that, unlike the ZSB constraints, conditional moment constraints do not require that action space  $\mathcal{T}$  is discrete and finite. We can again combine these conditional moment constraints (CMC) with relaxed conditional f-constraint (6) to obtain

$$\mathcal{W}_{f_{t,x}}^{\text{CMC}} := \left\{ w \geq 0 : \begin{array}{l} \mathbb{E}_{X,U} [f_{T,X} (p_{\text{obs}}(T|X)w)] \leq \gamma, \\ \mathbb{E}_{\text{obs}} [w|t,x] \cdot p_{\text{obs}}(t|x) = 1 \\ \text{for any } t \in \mathcal{T} \text{ and } x \in \mathcal{X} \end{array} \right\}$$

and its corresponding lower bound

$$V_{\text{inf}}^{\text{CMC}}(\pi) := \inf_{w \in \mathcal{W}_{f_{t,x}}^{\text{CMC}}} \mathbb{E}_{\text{obs}} [w(Y,T,X)\pi(T|X)Y]. \quad (11)$$

Indeed, it can be shown that the conditional moment constraints are strictly sharper than the ZSB constraints as discussed in the supplementary material. Therefore, one can naturally obtain  $\mathcal{W}_{f,t,x} \subseteq \mathcal{W}_{f,t,x}^{\text{CMC}} \subseteq \mathcal{W}_{f,t,x}^{\text{ZSB}}$  and  $V_{\text{inf}} \geq V_{\text{inf}}^{\text{CMC}} \geq V_{\text{inf}}^{\text{ZSB}}$ .

### 3 KERNEL CONDITIONAL MOMENT CONSTRAINTS

In this section, we introduce an empirical approximation of the conditional moment constraints using the kernel method (Schölkopf et al., 2002). The key idea is to approximate conditional moment  $\mathbb{E}_{\text{obs}}[w(Y, T, X)p_{\text{obs}}(T|X)|T = t, X = x]$  using the kernel ridge regression. By constraining the estimated conditional moment to be close to 1, we impose the conditional moment constraints to the empirical weight  $\mathbf{w} = (w_1, \dots, w_n)^T = (w(Y_1, T_1, X_1), \dots, w(Y_n, T_n, X_n))^T$ . In the following, we introduce three types of kernel conditional moment constraints (KCMC), namely, Gaussian process constraints, low-rank Gaussian process constraints, and low-rank hard constraints. These kernel conditional moment constraints are all convex constraints, and they enable us to define a tractable uncertainty set and an associated estimator of lower bound as

$$\hat{\mathcal{W}}_{f,t,x}^{\text{KCMC}} = \left\{ \mathbf{w} \geq 0 : \hat{\mathbb{E}}_n [f_{T,X}(p_{\text{obs}}(T|X)\mathbf{w})] \leq \gamma, \right. \\ \left. w(y, t, x) \text{ satisfies the KCMC} \right\}$$

and

$$\hat{V}_{\text{inf}}^{\text{KCMC}} := \min_{\mathbf{w} \in \hat{\mathcal{W}}_{f,t,x}^{\text{KCMC}}} \hat{\mathbb{E}}_n [w(Y, T, X)\pi(T|X)Y].$$

#### 3.1 Gaussian Process Constraints

In this subsection, we derive our first kernel conditional moment constraints, which we call the Gaussian process constraints. We begin by formally formulating the idea of using the kernel ridge regression for the conditional moment constraints and then motivate its interpretation as a Gaussian process to obtain reasonable kernel conditional moment constraints.

#### Estimation of Conditional Expectation by Kernel Ridge Regression

Let us introduce kernel  $k : (\mathcal{T} \times \mathcal{X}) \times (\mathcal{T} \times \mathcal{X}) \rightarrow \mathbb{R}$  with associated reproducing kernel Hilbert space (RKHS)  $\mathcal{H}$  of functions  $h : \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}$ , inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ , and norm  $\|\cdot\|_{\mathcal{H}}$ . Let us further introduce re-parametrization

$$e(y, t, x) := p_{\text{obs}}(t|x)w(y, t, x) - 1,$$

so that conditional moment constraints (10) can be written as

$$\mathbb{E}_{\text{obs}}[e(Y, T, X)|T = t, X = x] = 0$$

for any  $t \in \mathcal{T}$  and  $x \in \mathcal{X}$ .

Then, using the kernel ridge regression, one can estimate conditional expectation  $g(t, x) := \mathbb{E}_{\text{obs}}[e(Y, T, X)|T = t, X = x]$  as

$$\hat{g} = \arg \min_{g \in \mathcal{H}} \hat{\mathbb{E}}_n |g(T, X) - e(Y, T, X)|^2 + \sigma^2 \|g\|_{\mathcal{H}}^2.$$

for some  $\sigma^2 > 0$ . The above problem yields an analytical solution, and we can get

$$\hat{g} = K(K + \sigma^2 I_n)^{-1} \mathbf{e}$$

for  $\hat{g} := (\hat{g}(T_1, X_1), \dots, \hat{g}(T_n, X_n))^T$  and  $\mathbf{e} := (e(Y_1, T_1, X_1), \dots, e(Y_n, T_n, X_n))^T$ . Here,  $K$  denotes the kernel matrix such that  $K_{i,j} = k((T_i, X_i), (T_j, X_j))$  and  $I_n$  is the identity matrix of order  $n$ . To impose the conditional moment constraints, we can consider the constraint  $\hat{g} \approx \mathbf{0}$ . Note that we cannot impose the exact equality, i.e.,  $K(K + \sigma^2 I_n)^{-1} \mathbf{e} = \mathbf{0}$ , as this leads to only solution  $\mathbf{w} = \mathbf{1}/p_{\text{obs},T|X}$ ,<sup>3</sup> whose resulting policy value estimator is exactly the confounded IPW estimator. Therefore, we need to find a reasonable way to impose this close-to-zero constraint.

#### Construction of Uncertainty Set With a Credible Set of Gaussian Process

Now, by interpreting the kernel ridge regression as the Gaussian processes regression (Rasmussen, 2003), we make an intuitive association of the close-to-zero constraint with the credible set in Bayesian statistics. Consider the following Gaussian process regression model (Rasmussen, 2003):

$$g \sim \mathcal{GP}(0, k(\cdot, \cdot)), \quad \mathbf{e} = \mathbf{g} + \varepsilon, \quad \varepsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_n),$$

where  $\mathbf{g} := (g(T_1, X_1), \dots, g(T_n, X_n))$ .  $\mathcal{GP}$  and  $\mathcal{N}$  indicate the Gaussian process and the multivariate normal distribution with the specified mean and covariance parameters. Under this model, the posterior of  $\mathbf{g}$  given  $\mathbf{e}$  is a multivariate normal distribution with mean and variance  $\mu_{g|\mathbf{e}} = K(K + \sigma^2 I_n)^{-1} \mathbf{e} = \hat{g}$  and  $\Sigma_{g|\mathbf{e}} = K - K(K + \sigma^2 I_n)^{-1} K$ . For the posterior of  $\mathbf{g}$ , we now define a  $(1 - \alpha)$  credible set. As the posterior of  $\mathbf{g}$  given  $\mathbf{e}$  is a multivariate normal distribution, we can take highest posterior density set

$$C_{g|\mathbf{e}} = \{\mathbf{g} : (\mathbf{g} - \mu_{g|\mathbf{e}})^T \Sigma_{g|\mathbf{e}}^{-1} (\mathbf{g} - \mu_{g|\mathbf{e}}) \leq \chi_n^2(1 - \alpha)\}. \quad (12)$$

as the credible set. Here,  $\chi_n^2(1 - \alpha)$  denotes the  $(1 - \alpha)$ -percentile of  $\chi^2$  distribution with degrees of freedom  $n$ . Now, as we want estimated conditional moment function  $g(t, x)$  to be close to zero, we can require that  $\mathbf{g} = \mathbf{0}$  to be included in the  $(1 - \alpha)$  credible set and obtain the Gaussian process kernel conditional moment constraints as follows:

$$\text{KCMC}_{\text{GP}} \stackrel{\text{def}}{\Leftrightarrow} \mathbf{0} \in C_{g|\mathbf{e}} \\ \Leftrightarrow \mathbf{e}^T M_{\text{GP}} \mathbf{e} \leq \chi_n^2(1 - \alpha)$$

<sup>3</sup>Here,  $\mathbf{p}_{\text{obs},T|X} := (p_{\text{obs}}(T_1|X_1), \dots, p_{\text{obs}}(T_n|X_n))^T$  and  $\mathbf{1} := (1, \dots, 1) \in \mathbb{R}^n$ . The division is taken element-wise.

where  $M_{\text{GP}} := (K + \sigma^2 I_n)^{-1} K (K - K(K + \sigma^2 I_n)^{-1} K)^{-1} K (K + \sigma^2 I_n)^{-1}$ . As  $\mathbf{e} = \mathbf{p}_{\text{obs}, T|X} \odot \mathbf{w} - \mathbf{1}$ ,<sup>4</sup> this is a quadratic constraint for  $\mathbf{w}$ , which makes it possible to compute the associated lower bound by solving convex programming.

### 3.2 Low-rank Gaussian Process Constraints

A practical downside of the above formulation is the linear growth of the number of constraints to sample size  $n$ . Additionally, the calculation of the matrix inverse takes  $\mathcal{O}(n^3)$ , which can be impractically slow for a large sample size. To mitigate these issues, we propose the use of low-rank approximation to the kernel matrix.

Let's consider spectral decomposition of the kernel matrix,  $K = V \Lambda V^T$  with orthonormal matrix  $V \in \mathbb{R}^{n \times n}$  and diagonal matrix  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n \times n}$  satisfying  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ . By truncating the spectrum after the first  $D$  dimensions, we approximate the kernel matrix as  $K \approx \tilde{K} := \tilde{V} \tilde{\Lambda} \tilde{V}^T$ , where  $\tilde{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_D) \in \mathbb{R}^{D \times D}$  and  $\tilde{V} \in \mathbb{R}^{n \times D}$  is the first  $D$  columns of matrix  $V$ . Then, we can substitute  $\tilde{K}$  in place of  $K$  to obtain the approximate posterior of  $\mathbf{g}$ , which is a multivariate normal distribution with mean and variance  $\tilde{\mu}_{\mathbf{g}|e} = \tilde{K}(\tilde{K} + \sigma^2 I_n)^{-1} \mathbf{e}$  and  $\tilde{\Sigma}_{\mathbf{g}|e} = \tilde{K} - \tilde{K}(\tilde{K} + \sigma^2 I_n)^{-1} \tilde{K}$ . Thus, we can analogously define the credible set and the kernel conditional moment constraints as

$$\tilde{C}_{\mathbf{g}|e} = \{\mathbf{g} : (\mathbf{g} - \tilde{\mu}_{\mathbf{g}|e})^T \tilde{\Sigma}_{\mathbf{g}|e}^{-1} (\mathbf{g} - \tilde{\mu}_{\mathbf{g}|e}) \leq \chi_D^2(1 - \alpha)\}. \quad (13)$$

and

$$\begin{aligned} \text{KCMC}_{\text{low-rank GP}} &\stackrel{\text{def}}{\Leftrightarrow} \mathbf{0} \in \tilde{C}_{\mathbf{g}|e} \\ &\Leftrightarrow \mathbf{e}^T M_{\text{low-rank GP}} \mathbf{e} \leq \chi_D^2(1 - \alpha) \end{aligned} \quad (14)$$

where  $M_{\text{low-rank GP}} := (\tilde{K} + \sigma^2 I_n)^{-1} \tilde{K} (\tilde{K} - \tilde{K}(\tilde{K} + \sigma^2 I_n)^{-1} \tilde{K})^{-1} \tilde{K} (\tilde{K} + \sigma^2 I_n)^{-1}$ .

One big difference of the credible set of original Gaussian process (12) and low-rank Gaussian process (13) is the degree of freedom of the  $\chi^2$  distribution. We will discuss the reason for it in the supplementary material, but this is essentially due to the fact that the distribution of the  $\mathbf{g}$  is restricted to the  $D$ -dimensional subspace spanned by columns of  $\tilde{V}$ .

### 3.3 Low-rank Hard Constraints

Indeed, when we use the low-rank approximation, it is possible to impose the hard constraint to the solution of the low-rank kernel ridge regression as  $\hat{\mathbf{g}} = \tilde{K}(\tilde{K} + \sigma^2 I_n)^{-1} \mathbf{e} = \mathbf{0}$  without suffering from the issue of feasibility set reducing to singleton  $\{\mathbf{1}/\mathbf{p}_{\text{obs}, T|X}\}$ .

Using the interpretation of the spectral decomposition as the kernel principal component analysis (PCA) (Schölkopf

et al., 1997), the above hard constraints can be reformulated as the following empirical orthogonality condition

$$\hat{\mathbb{E}}_n[e(Y, T, X) \boldsymbol{\varphi}^{\text{KPCA}}(T, X)] = \mathbf{0}$$

of error  $e(Y, T, X)$  and kernel principal components  $\boldsymbol{\varphi}^{\text{KPCA}}(t, x) := (\varphi_1^{\text{KPCA}}, \dots, \varphi_D^{\text{KPCA}})^T$  obtained of the kernel PCA applied to  $\{T_i, X_i\}_{i=1}^n$ . This condition is obviously a relaxation of the conditional moment constraints (10), as they imply orthogonality condition  $\mathbb{E}_{\text{obs}}[e(Y, T, X) \boldsymbol{\psi}(T, X)] = \mathbf{0}$  for any  $\boldsymbol{\psi}(t, x)$ . By slightly generalizing the above condition, we introduce the following low-rank hard constraints.

$$\text{KCMC}_{\text{low-rank orth}} \stackrel{\text{def}}{\Leftrightarrow} \hat{\mathbb{E}}_n[e(Y, T, X) \boldsymbol{\psi}(T, X)] = \mathbf{0}, \quad (15)$$

for orthogonal function class  $\boldsymbol{\psi}(t, x) := (\psi_1, \dots, \psi_D)^T$ .

In our theoretical analysis, constraints (15) provide the most suitable estimator for our studies. This is because choosing  $\{\varphi_d\}_{d=1}^D$  independently from the samples enables us to decouple the discussion on the goodness of the constraints and the goodness of the policy value estimator.

In practice, these low-rank estimators of the lower bound have a trade-off between sharpness and credibility. Low-rank Gaussian process constraints (14) only impose soft quadratic constraints and can sometimes produce a too pessimistic lower bound. On the other hand, low-rank hard constraints (15) produce a tighter lower bound but require careful selection of the number of constraints, as excessively strong constraints would lead to a too optimistic estimate.

## 4 THEORETICAL ANALYSIS

In this section, we study the property of the kernel conditional moment constraints for confounding robust inference. For the convenience of the analysis, we only consider low-rank orthogonality condition (15) as the kernel conditional moment constraints in this section. Thus, we will focus on the property of the following population lower bound

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

for

$$\mathcal{W}_{f_{t,x}}^{\text{KCMC}} = \left\{ w \geq 0 : \begin{aligned} &\mathbb{E}_{\text{obs}}[f_{T,X}(w p_{\text{obs}}(T|X))] \leq \gamma, \\ &\mathbb{E}_{\text{obs}}[(w p_{\text{obs}}(T|X) - 1) \boldsymbol{\psi}] = \mathbf{0} \end{aligned} \right\}$$

and its empirical version

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

for

$$\hat{\mathcal{W}}_{f_{t,x}}^{\text{KCMC}} = \left\{ w \geq 0 : \begin{aligned} &\hat{\mathbb{E}}_n[f_{T,X}(w p_{\text{obs}}(T|X))] \leq \gamma, \\ &\hat{\mathbb{E}}_n[(w p_{\text{obs}}(T|X) - 1) \boldsymbol{\psi}] = \mathbf{0} \end{aligned} \right\}.$$

<sup>4</sup>The element-wise product operator is denoted by  $\odot$ .

In the theoretical analysis, we characterize the properties of our estimator with the dual solution of our original problem. We first analyze specification error of the kernel conditional moment constraints  $|V_{\text{inf}}^{\text{KCMC}} - \hat{V}_{\text{inf}}^{\text{KCMC}}|$  and provide a condition on orthogonal function class  $\{\psi_d(t, x)\}_{d=1}^D$  under which the specification error becomes zero. Additionally, we study empirical estimator  $\hat{V}_{\text{inf}}^{\text{KCMC}}$  and prove consistency guarantees for policy evaluation and learning.

Due to the space limitation, we will defer the proofs and the precise assumptions for the following statements to the supplementary material.

#### 4.1 Specification Error

First, we present the condition under which the specification error of estimator  $V_{\text{inf}}^{\text{KCMC}}(\pi)$  becomes zero for policy  $\pi$  such that  $|V_{\text{inf}}^{\text{KCMC}}(\pi) - \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi)| = 0$ .

**Theorem 1 (No specification error)** *Let  $\eta_{\text{CMC}}^* : \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}$  be the solution to the dual problem of (11). Then  $V_{\text{inf}}^{\text{KCMC}}(\pi)$  has zero specification error if*

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

Using this lemma, it is possible to prove that the previous sharp estimator for box constraints by Dorn and Guo (2022) is a special case of estimator that uses our kernel conditional moment constraints. They identified the analytical form for optimal orthogonal function class  $\{\psi_1\} = \{\eta_{\text{CMC}}^*\}$  to derive a one-dimensional linear constraint to impose the conditional moment constraints, which are originally infinite-dimensional. More details on their estimator are discussed in the supplementary material.

#### 4.2 Consistency of Policy Evaluation and Learning

Now, we study empirical estimator  $\hat{V}_{\text{inf}}^{\text{KCMC}}$  and provide consistency guarantees for policy evaluation and learning in the case of finite-dimensional concave policy class. We prove both consistency results by a reduction of our problem to the M-estimation (Van de Geer, 2000) using the dual formulation.

Let us define  $L_{\theta, \pi} : \mathcal{Y} \times \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}$  as the loss function of the dual objective of (16) so that the dual problem becomes  $\max_{\theta \in \Theta} \mathbb{E}[-L_{\theta, \pi}(Y, T, X)]$  for dual parameter  $\theta \in \Theta$ . Additionally, let us introduce the following assumption:

**Assumption 1 (Regularity of loss function)** *Loss function  $\ell_{\theta} : \mathcal{Y} \times \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}$  satisfies 1)  $\theta \mapsto \ell_{\theta}$  is continuous, 2)  $\mathbb{E}|\ell_{\theta}| < \infty$  for any  $\theta \in \Theta$ , 3)  $\theta_0 \in \arg \min_{\theta \in \Theta} \mathbb{E}[\ell_{\theta}]$  is unique, and 4)  $\mathbb{E}[G_{\varepsilon}] < \infty$  for  $G_{\varepsilon} := \sup_{\theta \in \Theta: \|\theta - \theta_0\| \leq \varepsilon} |\ell_{\theta}|$  for some  $\varepsilon > 0$ .*

With this assumption, we can immediately show the consistency guarantee for policy evaluation:

**Theorem 2 (Consistency of policy evaluation)** *For fixed policy  $\pi$ , if  $\ell_{\theta} := L_{\theta, \pi}$  satisfies Assumption 1, then we have  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi) \xrightarrow{P} V_{\text{inf}}^{\text{KCMC}}(\pi)$ .*

The above theorem can be extended to policy learning, by considering joint parameter space  $\Theta' := \Theta \times \mathcal{B}$  of  $\theta' := (\theta, \beta)$  for policy class  $\{\pi_{\beta}(t|x) : \beta \in \mathcal{B}\}$  and introducing joint loss function  $L'_{\theta'} := L_{\theta, \pi_{\beta}}$  so that  $\max_{\beta \in \mathcal{B}} V_{\text{inf}}^{\text{KCMC}}(\pi_{\beta}) = \max_{\theta' \in \Theta'} \mathbb{E}[-L'_{\theta'}(Y, T, X)]$ .

**Theorem 3 (Consistency of concave policy learning)**

*Assume the policy class is concave so that  $\beta \mapsto \pi_{\beta}(t|x)y$  is concave. Define  $\beta_0 \in \arg \max_{\beta \in \mathcal{B}} V_{\text{inf}}^{\text{KCMC}}(\pi_{\beta})$  and its estimator  $\hat{\beta} \in \arg \max_{\beta \in \mathcal{B}} \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\beta})$ . If  $L'_{\theta'}$  satisfies Assumption 1, then, we have  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) \xrightarrow{P} V_{\text{inf}}^{\text{KCMC}}(\pi_{\beta_0})$ .*

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.

## 5 NUMERICAL EXPERIMENTS

In this section, we present numerical examples to compare our estimators with the existing estimators. In addition to the standard policy evaluation, we also consider policy learning and the f-sensitivity model.

### 5.1 Experimental Settings

In the first three experiments, we use the synthetic data adapted from Kallus and Zhou (2018, 2021). We repeat the experiment 10 times with different random seeds and report the mean and one standard deviation range by a line and a band around it. The last experiment uses subsamples of data from the 1966-1981 National Longitudinal Survey (NLS) of Older and Young Men, which was also used in Dorn and Guo (2022).

In the experiments, four types of estimators are compared. As the baseline, we consider the conventionally used ZSB estimator which solves the empirical version of (9). To this baseline, we compare the proposed estimators based on two types of kernel conditional moment constraints (KCMC), which are (14) and (15). We call them the low-rank GP KCMC and the low-rank hard KCMC, respectively. For low-rank hard KCMC, the orthogonal function class was chosen by the kernel PCA. Lastly, as a reference, we include the quantile balancing (QB) estimator by Dorn et al. (2021), which is a special case of low-rank hard KCMC. More details on the experimental settings and additional results can be found in the supplementary material.<sup>5</sup>

<sup>5</sup>The code can be found at <https://github.com/kstoneriv3/confounding-robust-inference>.

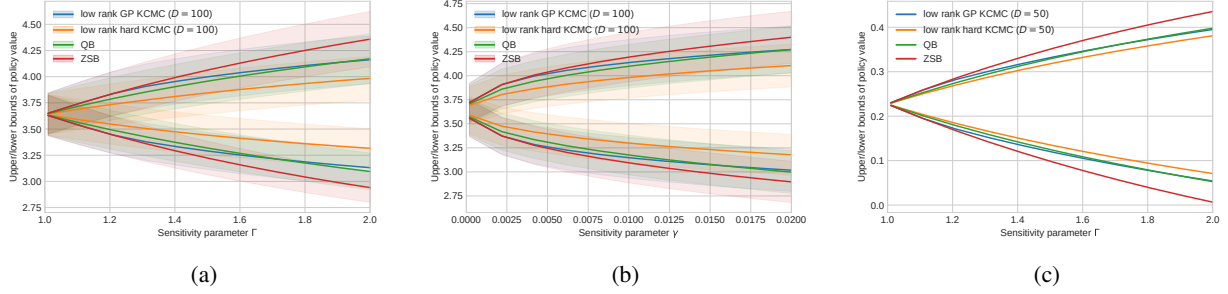


Figure 1: Estimated upper and lower bounds using different types of estimators for sensitivity analysis: (a) Tan’s marginal sensitivity model for policy value on synthetic data; (b) The KL sensitivity model for policy value on synthetic data; (c) Tan’s marginal sensitivity model ( $\Gamma = 1.5$ ) for average treatment effect on NLS data.

## 5.2 Policy Evaluation

Figure 1a compares the tightness of the bounds obtained by different estimators in policy evaluation. Clearly, the sharper estimators (KCMC and QB) are producing tighter bounds than the ZSB estimator. Here, we can see that the low-rank hard KCMC’s bounds are much tighter than those of the other sharp estimators. This exemplifies the aforementioned trade-off between the sharpness and the credibility of bounds obtained by the soft and hard KCMC.

## 5.3 Extension to f-divergence Sensitivity Model

Next, to illustrate application of the f-sensitivity models (3), we present an example of the KL-sensitivity model in Figure 1b. We can see that the KL-sensitivity model can provide continuous control of the level of confounding by the sensitivity parameter, similarly to Tan’s marginal sensitivity model.

## 5.4 Extension to Policy Learning

Figure 2 shows the learning curves during the max-min policy optimization with Tan’s marginal sensitivity model. Though order  $\hat{V}_{\text{inf}}^{\text{KCMC}} \geq \hat{V}_{\text{inf}}^{\text{ZSB}}$  is still maintained, the KCMC lower bounds estimated from training data are actually higher than the ground truth. This can be interpreted as an overfitting phenomenon in the joint maximization of learning and the inner dual problem. Since  $V_{\text{inf}}^{\text{CMC}}$  estimated by test data is still lower than the ground truth, improvement of policy learning may be possible with more careful control of test errors such as cross validation.

## 5.5 Treatment Effect Estimation on NLS Data

Lastly, in Figure 1c, we show the upper and lower bounds of the average treatment effect ( $\mathbb{E}[Y|T = 1] - \mathbb{E}[Y|T = 0]$ ) of union membership ( $T$ ) on log wages ( $Y$ ) estimated from the NLS dataset. Similarly to the previous examples, the low-rank GP KCMC and quantile balancing estimates are

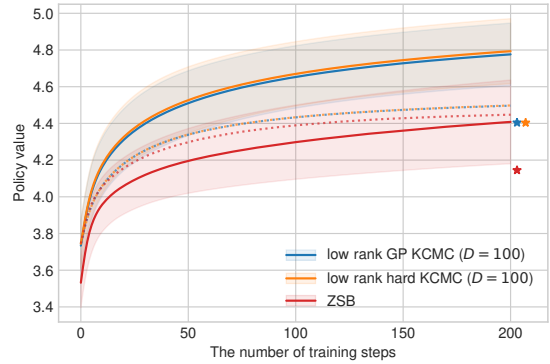


Figure 2: The value of  $\hat{V}_{\text{inf}}$  on training data during the policy learning. The dotted lines represent ground truth policy values  $V$ . The star symbols ( $\star$ ) at the end of the learning curves indicate the average lower bound  $V_{\text{inf}}^{\text{CMC}}$  of the learned policy.<sup>6</sup>

very close while the ZSB and the low-rank hard KCMC produce looser and (possibly overly) tighter bounds.

## 6 CONCLUSION

In this paper, we proposed kernel approximation of the conditional moment constraints to achieve sharp and general sensitivity analysis. We theoretically studied the property of the kernel conditional moment constraints and established consistency guarantees for policy evaluation and learning. We also confirmed the effectiveness of our approach empirically, with numerical examples covering various types of problems in the sensitivity analysis.

## Acknowledgements

K.I. was supported by the Heiwa Nakajima Foundation. N.H. was supported by ETH Research Grant and Swiss National Science Foundation. The authors acknowledge Takafumi Kanamori (Tokyo Institute of Technology) for pointing out a critical technical error in an early version of the draft.

<sup>6</sup>Ground truth  $V$  was estimated by unconfounded Monte Carlo simulation of the true data-generating process. Lower bound  $V_{\text{inf}}^{\text{CMC}}$  was estimated by the low-rank hard KCMC estimators on test data.



## References

- ApS, M. (2019). Mosek optimization suite.
- Aubin-Frankowski, P.-C. and Szabó, Z. (2020). Hard shape-constrained kernel machines. *Advances in Neural Information Processing Systems*, 33:384–395.
- Boyd, S., Boyd, S. P., and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.
- Diamond, S. and Boyd, S. (2016). CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83):1–5.
- Domahidi, A., Chu, E., and Boyd, S. (2013). ECOS: An SOCP solver for embedded systems. In *European Control Conference (ECC)*, pages 3071–3076.
- Dorn, J. and Guo, K. (2022). Sharp sensitivity analysis for inverse propensity weighting via quantile balancing. *Journal of the American Statistical Association*.
- Dorn, J., Guo, K., and Kallus, N. (2021). Doubly-valid/doubly-sharp sensitivity analysis for causal inference with unmeasured confounding. *arXiv preprint arXiv:2112.11449*.
- Hirano, K. and Imbens, G. W. (2001). 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.
- Hirano, K., Imbens, G. W., and Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189.
- Jin, Y., Ren, Z., and Zhou, Z. (2022). Sensitivity analysis under the  $f$ -sensitivity models: Definition, estimation and inference. *arXiv preprint arXiv:2203.04373*.
- Kallus, N. (2018). Balanced policy evaluation and learning. *Advances in neural information processing systems*, 31.
- Kallus, N., Mao, X., and Uehara, M. (2021). Causal inference under unmeasured confounding with negative controls: A minimax learning approach. *arXiv preprint arXiv:2103.14029*.
- Kallus, N. and Zhou, A. (2018). Confounding-robust policy improvement. *Advances in neural information processing systems*, 31.
- Kallus, N. and Zhou, A. (2021). Minimax-optimal policy learning under unobserved confounding. *Management Science*, 67(5):2870–2890.
- Kremer, H., Zhu, J.-J., Muandet, K., and Schölkopf, B. (2022). Functional generalized empirical likelihood estimation for conditional moment restrictions. In *International Conference on Machine Learning*, pages 11665–11682. PMLR.
- Liu, W., Kuramoto, S. J., and Stuart, E. A. (2013). An introduction to sensitivity analysis for unobserved confounding in nonexperimental prevention research. *Prevention science*, 14(6):570–580.
- Mastouri, A., Zhu, Y., Gultchin, L., Korba, A., Silva, R., Kusner, M., Gretton, A., and Muandet, K. (2021). Proximal causal learning with kernels: Two-stage estimation and moment restriction. In *International Conference on Machine Learning*, pages 7512–7523. PMLR.
- Muandet, K., Jitkrittum, W., and Kübler, J. (2020). Kernel conditional moment test via maximum moment restriction. In *Conference on Uncertainty in Artificial Intelligence*, pages 41–50. PMLR.
- Pérez-Suay, A., Laparra, V., Mateo-García, G., Muñoz-Marí, J., Gómez-Chova, L., and Camps-Valls, G. (2017). Fair kernel learning. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 339–355. Springer.
- Rasmussen, C. E. (2003). Gaussian processes in machine learning. In *Summer school on machine learning*, pages 63–71. Springer.
- Rosenbaum, P. R. (2002). Overt bias in observational studies. In *Observational studies*, pages 71–104. Springer.
- Rosenbaum, P. R., Rosenbaum, P., and Briskman (2010). *Design of observational studies*, volume 10. Springer.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688.
- Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469):322–331.
- Schölkopf, B., Smola, A., and Müller, K.-R. (1997). Kernel principal component analysis. In *International conference on artificial neural networks*, pages 583–588. Springer.
- Schölkopf, B., Smola, A. J., Bach, F., et al. (2002). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press.
- Singh, R. (2020). Kernel methods for unobserved confounding: Negative controls, proxies, and instruments. *arXiv preprint arXiv:2012.10315*.
- Singh, R., Sahani, M., and Gretton, A. (2019). Kernel instrumental variable regression. *Advances in Neural Information Processing Systems*, 32.
- Staib, M. and Jegelka, S. (2019). Distributionally robust optimization and generalization in kernel methods. *Advances in Neural Information Processing Systems*, 32.
- Tan, Z. (2006). A distributional approach for causal inference using propensity scores. *Journal of the American Statistical Association*, 101(476):1619–1637.
- Van de Geer, S. (2000). *Empirical Processes in M-estimation*, volume 6. Cambridge university press. We referenced the author’s lecture note based on the book, available online at <https://stat.ethz.ch/~geer/empirical-processes.pdf>.

- Zhao, Q., Small, D. S., and Bhattacharya, B. B. (2019). 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.
- Zhu, J.-J., Jitkrittum, W., Diehl, M., and Schölkopf, B. (2020). 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.

## A Derivation of ZSB Constraints and Conditional Moment Constraints

Here, we discuss the two ways to relax the distributional constraint of the original uncertainty set for base policy  $\pi_{\text{base}}$ . The condition that  $\pi_{\text{base}}(t|x, u)$  is a proper distribution is equivalent to

$$\sum_{\mathcal{T}} \pi_{\text{base}}(t|x, u) dt' = 1 \text{ and } \pi_{\text{base}}(t|x, u) \geq 0 \text{ for any } t \in \mathcal{T}, x \in \mathcal{X}, \text{ and } u \in \mathcal{U}. \quad (17)$$

These constraints have traditionally been relaxed to the ZSB constraints, which are used in the well-known Hájek estimator (Zhao et al., 2019; Kallus and Zhou, 2018, 2021). However, there is also a tighter relaxation called conditional moment constraints, which we employ in our work. In the following, we present these two types of relaxation.

### A.1 ZSB Constraint

When action space  $\mathcal{T}$  is discrete and finite, a well-known relaxation of (17) is

$$\mathbb{E}_{\text{obs}}[\mathbb{1}_{T=t} w(Y, T, X)] = 1 \text{ for any } t \in \mathcal{T} \quad (18)$$

and

$$w(y, t, x) \geq 0 \text{ for any } y \in \mathcal{Y}, t \in \mathcal{T}, \text{ and } x \in \mathcal{X}, \quad (19)$$

where  $\mathbb{1}$  denotes the indicator function for event  $A$ . Following the naming convention in Dorn and Guo (2022), we will call this constraint the ZSB constraint after the authors of Zhao et al. (2019). Condition (18) can be obtained as

$$\begin{aligned} \mathbb{E}_{\text{obs}}[w(Y, T, X) \mathbb{1}_{T=t}] &= \mathbb{E}_{T \sim \pi_{\text{base}}(\cdot|X, U)} \left[ \frac{\mathbb{1}_{T=t}}{\pi_{\text{base}}(T|X, U)} \right] \\ &= \mathbb{E}_{X, U} \left[ \int_{\mathbb{R}} \sum_{\mathcal{T}} \left( \frac{\mathbb{1}_{t'=t}}{\pi_{\text{base}}(t'|X, U)} \right) p(y|t', X, U) \pi_{\text{base}}(t'|X, U) dy dt' \right] \\ &= \mathbb{E}_{X, U} \left[ \sum_{\mathcal{T}} \mathbb{1}_{t'=t} \int_{\mathbb{R}} p(y|t', X, U) dt' dy \right] = 1, \end{aligned}$$

and the non-negativity condition is trivial from the definition of  $w(y, t, x)$ . Combining the above with the relaxation of the conditional f-constraint, the following uncertainty set with the ZSB constraint can be defined:

$$\mathcal{W}_{f_{t,x}}^{\text{ZSB}} := \left\{ w(y, t, x) \geq 0 : \begin{array}{l} \mathbb{E}_{X, U} [f_{T, X} (p_{\text{obs}}(T|X) w(Y, T, X))] \leq \gamma \\ \text{and} \\ \mathbb{E}_{\text{obs}}[\mathbb{1}_{T=t} w(Y, T, X)] = 1 \text{ for any } t \in \mathcal{T} \end{array} \right\}.$$

This uncertainty set has been traditionally adopted by many works such as Tan (2006); Zhao et al. (2019); Kallus and Zhou (2018, 2021). For this uncertainty set, associated lower bound

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

can be consistently approximated straightforwardly. By approximating the expectations by empirical average, we get a linear program with parameter  $\mathbf{w} = (w_1, \dots, w_n)^T = (w(Y_1, T_1, X_1), \dots, w(Y_n, T_n, X_n))^T$ ,

$$\hat{V}_{\text{inf}}^{\text{ZSB}}(\pi) := \min_{\mathbf{w} \in \hat{\mathcal{W}}_{f_{t,x}}^{\text{ZSB}}} \hat{\mathbb{E}}_n [w(Y, T, X) \pi(T|X) Y] \quad (20)$$

where

$$\hat{\mathcal{W}}_{f_{t,x}}^{\text{ZSB}} := \left\{ \mathbf{w} \geq \mathbf{0} : \begin{array}{l} \hat{\mathbb{E}}_n [f_{T, X} (p_{\text{obs}}(T|X) w(Y, T, X))] \leq \gamma \\ \text{and} \\ \hat{\mathbb{E}}_n [\mathbb{1}_{T=t} w(Y, T, X)] = 1 \text{ for all } t \in \mathcal{T} \end{array} \right\}.$$

Here we should note that (20) is not exactly the estimator used in Zhao et al. (2019) and its recent extensions. Instead, they solved  $\min_{\mathbf{w} \geq \mathbf{0}} \frac{\hat{\mathbb{E}}_n [w(Y, T, X) \pi(T|X) Y]}{\hat{\mathbb{E}}_n [w(Y, T, X)]}$  such that  $\hat{\mathbb{E}}_n [f_{T, X} (p_{\text{obs}}(T|X) w(Y, T, X))] \leq \gamma$ , by the linear fractional programming.

## A.2 Conditional Moment Constraints

Now, we introduce the sharper constraints that we leverage in our work. We relax constraint (17) as conditional moment constraints

$$\mathbb{E}_{\text{obs}}[w(Y, T, X)|T = t, X = x] \cdot p_{\text{obs}}(t|x) = 1 \text{ for any } t \in \mathcal{T} \text{ and } x \in \mathcal{X} \quad (21)$$

plus non-negativity constraint (19). For this relaxation, we do not require that action space  $\mathcal{T}$  is discrete and finite. We can check the validity of this relaxation from

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

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 model. We again combine these conditional moment constraints (CMC) with the relaxed conditional f-constraint to obtain

$$\mathcal{W}_{f_{t,x}}^{\text{CMC}} := \left\{ w(y, t, x) \geq 0 : \begin{array}{l} \mathbb{E}_{X, U} [f_{T, X}(p_{\text{obs}}(T|X)w(Y, T, X))] \leq \gamma \\ \text{and} \\ \mathbb{E}_{\text{obs}}[w(Y, T, X)|T = t, X = x] \cdot p_{\text{obs}}(t|x) = 1 \text{ for any } t \in \mathcal{T} \text{ and } x \in \mathcal{X} \end{array} \right\}$$

and its corresponding lower bound

$$V_{\text{inf}}^{\text{CMC}}(\pi) := \inf_{w \in \mathcal{W}_{f_{t,x}}^{\text{CMC}}} \mathbb{E}_{\text{obs}} [w(Y, T, X)\pi(T|X)Y]. \quad (22)$$

Naturally, for these uncertainty sets, 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  $\mathcal{W}_{f_{t,x}}^{\text{CMC}}$ . We can show the latter inclusion by taking the (conditional) expectation of conditional moment constraints (21) with respect to  $p_{\text{obs}}(x|t)$  to obtain ZSB constraint (18).

Here, we can also show that there exist cases where strict inclusion holds so that these sets are not equivalent. For the latter inclusion, it is trivial to show that strict inclusion holds when the conditional moment constraints are stronger than the ZSB constraint. For the former inclusion, we can construct the following toy example where there exists some  $w(y, t, x) \in \mathcal{W}_{f_{t,x}}^{\text{CMC}} \setminus \mathcal{W}_{f_{t,x}}$ .

**Example 1 (A non-realizable element in  $\mathcal{W}_{f_{t,x}}^{\text{CMC}}$ )** Let us assume that the context space is a singleton and the action space and the reward space are binary so that  $\mathcal{X} = \{x\}$  and  $\mathcal{T} = \mathcal{Y} = \{-1, +1\}$ . Then, observational distribution

$$p_{\text{obs}}(Y = \pm 1, T = \pm 1, X = x) = 1/4$$

and inverse probability weight

$$w(y, t, x) = \begin{cases} 3.1 & \text{if } y = -1 \\ 0.9 & \text{if } y = +1 \end{cases}$$

satisfy conditional moment constraints (21) as well as non-negativity constraints (19) since

$$\mathbb{E}_{\text{obs}}[w(Y, T, X)|T = +1, X = x] \cdot p_{\text{obs}}(T = +1|X = x) = \left(0.9 \cdot \frac{1}{2} + 3.1 \cdot \frac{1}{2}\right) \cdot \frac{1}{2} = 1$$

and

$$\mathbb{E}_{\text{obs}}[w(Y, T, X)|T = -1, X = x] \cdot p_{\text{obs}}(T = -1|X = x) = \left(0.9 \cdot \frac{1}{2} + 3.1 \cdot \frac{1}{2}\right) \cdot \frac{1}{2} = 1.$$

By considering some conditional  $f$ -constraint that contains the above parameter values, we can construct some uncertainty set  $\mathcal{W}_{f,t,x}^{\text{CMC}}$  that contains this  $w(y, t, x)$  given above observational distribution  $p_{\text{obs}}(y, t, x)$ . However, there exists no proper policy  $\pi_{\text{base}}(t|x, u)$  and underlying model  $p(y|t, x, u)$  and  $p(x, u)$  that satisfy both

$$p_{\text{obs}}(y, t, x) = \int_{\mathcal{U}} p(y|t, x, u) \pi_{\text{base}}(t|x, u) p(x, u) du$$

and

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

Indeed,

$$\begin{aligned} p_{\text{obs}}(Y = +1, t, x) &= \int_{\mathcal{U}} p(y|t, x, u) \pi_{\text{base}}(t|x, u) p(x, u) du \\ &\leq \int_{\mathcal{U}} p(y|t, x, u) p(x, u) du \\ &= w(+1, t, x) \cdot p_{\text{obs}}(Y = +1, t, x) \end{aligned}$$

contradicts our assumptions on the model parameters  $p_{\text{obs}}(Y = +1, t, x) = 1/4$  and  $w(+1, t, x) = 0.9$ . In other words, such element  $w(y, x, t)$  in uncertainty set  $\mathcal{W}_{f,t,x}^{\text{CMC}}$  is not realizable, and thus, the strict inclusion holds for  $\mathcal{W}_{f,t,x} \subseteq \mathcal{W}_{f,t,x}^{\text{CMC}}$  in this case. This example exploits the too much flexibility of the conditional  $f$ -constraint by taking the unrealizable case where  $w(+1, t, x) < 1$ . In the case of discrete action space, we know that  $w \geq 1$  because the inverse probability of discrete action is always no less than 1.

Having shown the inclusion relations of the uncertainty sets, we can discuss the relations among the lower bounds. Assuming that the true base policy is contained in the original sensitivity model, we have  $V \geq V_{\text{inf}} \geq V_{\text{inf}}^{\text{CMC}} \geq V_{\text{inf}}^{\text{ZSB}}$ . One important question to ask here is under what kind of constraints the second equality holds so that the lower bound of the conditional moment constraints is tight. Surprisingly, recent work by Dorn and Guo (2022) showed that equality  $V_{\text{inf}} = V_{\text{inf}}^{\text{CMC}}$  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 such as the evaluation of general policy, non-binary action spaces, and general box and  $f$ -divergence constraints. Clearly, Example 1 shows that the same equality does not always hold for any box constraints, as it can be the minimizer for some box constraints and some policies. However, the question of which constraint class yields a tight bound with a realizable minimizer is an open question.

## B Alternative $f$ -sensitivity Models by Jin et al. (2022)

Here, we describe the difference between our  $f$ -sensitivity models and the  $f$ -sensitivity models proposed by Jin et al. (2022) that also uses uncertainty sets defined with the  $f$ -divergence. They proposed a similar uncertainty set that relaxes the condition

$$D_f [p(Y(1)|X, T = 1) || p(Y(1)|X, T = 0)] \leq \gamma,$$

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  $\gamma$ .

In terms of the 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 way of modeling is different from the one that Tan (2006) and its extension is based on.

## C Quantile Balancing Estimator by Dorn and Guo (2022)

In this section, we discuss the recently proposed tractable estimators for lower bound  $V_{\text{inf}}^{\text{CMC}}$  by Dorn and Guo (2022). As the optimization for  $V_{\text{inf}}^{\text{CMC}}$  involves infinite dimensional constraints for all  $t \in \mathcal{T}$  and  $x \in \mathcal{X}$ , we cannot apply the method analogous to (20) to obtain the empirical version of the lower bound. They studied box constraints  $a_w(t, x) \leq w(y, t, x) \leq b_w(t, x)$  in the case of discrete finite action space and proposed the first tractable method to impose such constraints. They showed that we can solve (22) in the case of box constraints as

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

subject to

$$\mathbb{E}_{\text{obs}}[w(Y, T, X)\pi(T|X)Q(T, X)] = \mathbb{E}_{\text{obs}} \left[ \left( \frac{\pi(T|X)}{p_{\text{obs}}(T|X)} \right) Q(T, X) \right], \quad (23)$$

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 the marginal sensitivity model by Tan (2006), the expression for  $\tau(t, x)$  can be simplified as  $\tau(x) = \frac{1}{1+\Gamma}$ . They leveraged the above characterization for  $V_{\text{inf}}^{\text{CMC}}$  by plugging in estimate  $\hat{Q}(t, x)$  of the conditional quantile function and empirically approximating the expectation to obtain a tractable linear programming problem for the quantile balancing (QB) estimator,

$$\hat{V}_{\text{inf, box}}^{\text{QB}} = \min_{a_w(t, x) \leq w_i \leq b_w(t, x)} \hat{\mathbb{E}}_n[w(Y, T, X)\pi(T|X)Y]$$

subject to

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

Indeed, we will show later that this quantile balancing estimator is a special case of our estimator where we have a nearly optimal choice of orthogonal function class  $\{\psi_1\}$  where  $\psi_1(x, t) = \left( \frac{\pi(t|x)}{p_{\text{obs}}(t|x)} \right) \hat{Q}(t, x)$ . Therefore, the KCMC estimator is guaranteed to be no looser than QB estimator when  $\left( \frac{\pi(t|x)}{p_{\text{obs}}(t|x)} \right) \hat{Q}(t, x) \in \text{span}(\{\psi_d(t, x)\}_{d=1}^D)$ . To put it in another way, if we estimate the quantile by linear quantile regression with feature vectors  $\left\{ \left( \frac{p_{\text{obs}}(t|x)}{\pi(t|x)} \right) \psi_d(x, t) \right\}_{d=1}^D$ , the QB estimator is no tighter than the KCMC estimator.

As our estimator generalizes the previous work, our estimator overcomes some drawbacks of the quantile balancing estimators. For example, the quantile balancing estimator cannot handle policy learning and the f-divergence constraint. Policy learning is 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 box constraints and does not have a proper extension to the f-sensitivity model. 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.

## D More Detail on Low-rank Gaussian Process Constraints

Here, we provide more discussion on the derivation and interpretations of the low-rank Gaussian process kernel conditional moment constraints.

### D.1 Derivation

Let's consider spectral decomposition of the kernel matrix  $K = V\Lambda V^T$  with orthonormal matrix  $V \in \mathbb{R}^{n \times n}$  and diagonal matrix  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n \times n}$  satisfying  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ . By truncating the spectrum after the first

$D$  dimensions, we approximate the kernel matrix as  $K \approx \tilde{K} := \tilde{V}\tilde{\Lambda}\tilde{V}^T$ , where  $\tilde{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_D) \in \mathbb{R}^{D \times D}$  and  $\tilde{V} \in \mathbb{R}^{n \times D}$  is the first  $D$  columns of matrix  $V$ . Then, we can obtain the low-rank version of the Gaussian process regression model as

$$\begin{aligned} \mathbf{z} &\sim \mathcal{N}(\mathbf{0}, \tilde{\Lambda}) \\ \mathbf{g} &= \tilde{V}\mathbf{z}, \\ \mathbf{e} &= \mathbf{g} + \varepsilon, \\ \varepsilon &\sim \mathcal{N}(\mathbf{0}, \sigma^2 I_n). \end{aligned}$$

Here, we can verify that this model is a valid low-rank approximation, by checking that the prior mean and covariance of  $\mathbf{g}$  are  $\mathbf{0}$  and  $\tilde{K}$  respectively.

For this low-rank model, the posterior distribution of  $\mathbf{g}$  given  $\mathbf{e}$  is again a multivariate normal distribution with mean and variance  $\tilde{\mu}_{\mathbf{g}|\mathbf{e}} = \tilde{K}(\tilde{K} + \sigma^2 I_n)^{-1}\mathbf{e}$  and  $\tilde{\Sigma}_{\mathbf{g}|\mathbf{e}} = \tilde{K} - \tilde{K}(\tilde{K} + \sigma^2 I_n)^{-1}\tilde{K}$ . Therefore, we can analogously define the credible set and the kernel conditional moment constraints as

$$\tilde{C}_{\mathbf{g}|\mathbf{e}} = \{\mathbf{g} : (\mathbf{g} - \tilde{\mu}_{\mathbf{g}|\mathbf{e}})^T \tilde{\Sigma}_{\mathbf{g}|\mathbf{e}}^{-1} (\mathbf{g} - \tilde{\mu}_{\mathbf{g}|\mathbf{e}}) \leq \chi_D^2(1 - \alpha)\}.$$

and

$$\begin{aligned} \text{KCMC}_{\text{low-rank GP}} &\stackrel{\text{def}}{\Leftrightarrow} \mathbf{0} \in \tilde{C}_{\mathbf{g}|\mathbf{e}} \\ &\Leftrightarrow (\mathbf{w} \odot \mathbf{p}_{\text{obs}, T|X} - 1)^T M_{\text{low-rank GP}} (\mathbf{w} \odot \mathbf{p}_{\text{obs}, T|X} - 1) \leq \chi_D^2(1 - \alpha) \end{aligned}$$

where  $M_{\text{low-rank GP}} := (\tilde{K} + \sigma^2 I_n)^{-1} \tilde{K} (\tilde{K} - \tilde{K}(\tilde{K} + \sigma^2 I_n)^{-1} \tilde{K})^{-1} \tilde{K} (\tilde{K} + \sigma^2 I_n)^{-1}$ .

A big difference between the credible set of the original Gaussian process and the low-rank Gaussian process is the degree of freedom of the  $\chi^2$  distribution. This is because the distribution of the  $\mathbf{g}$  is essentially restricted to the  $D$ -dimensional subspace spanned by columns of  $\tilde{V}$ . The condition that the credible set of  $\mathbf{g}$  includes  $\mathbf{0} \in \mathbb{R}^n$  is equivalent to the condition that  $\mathbf{0} \in \mathbb{R}^D$  is contained in the credible set of  $\mathbf{z}$ , whose posterior is  $D$ -dimensional multivariate normal distribution.

## D.2 A Spectral Interpretation of Low-rank and Full-rank Constraints

With the spectral decomposition, we can obtain more intuitive expressions for quadratic forms  $\mathbf{e}^T M_{\text{GP}} \mathbf{e}$  and  $\mathbf{e}^T M_{\text{low-rank GP}} \mathbf{e}$ , which are

$$\begin{aligned} \mathbf{e}^T M_{\text{GP}} \mathbf{e} &= \frac{1}{\sigma^2} (V^T \mathbf{e})^T \text{diag} \left( \frac{\lambda_1}{\lambda_1 + \sigma^2}, \dots, \frac{\lambda_n}{\lambda_n + \sigma^2} \right) (V^T \mathbf{e}), \\ \mathbf{e}^T M_{\text{low-rank GP}} \mathbf{e} &= \frac{1}{\sigma^2} (\tilde{V}^T \mathbf{e})^T \text{diag} \left( \frac{\lambda_1}{\lambda_1 + \sigma^2}, \dots, \frac{\lambda_D}{\lambda_D + \sigma^2} \right) (\tilde{V}^T \mathbf{e}). \end{aligned}$$

Now we can interpret the above quantities from the perspective of the kernel principal component analysis (KPCA) (Schölkopf et al., 1997). Let  $\varphi_1^{\text{KPCA}}, \dots, \varphi_n^{\text{KPCA}}$  be the empirically normalized principal components in RKHS corresponding to the empirical spectrum  $\lambda_1/n, \dots, \lambda_n/n$ , which satisfy  $\hat{\mathbb{E}}_n |\varphi_i^{\text{KPCA}}(Y, T, X)|^2 = 1$  for any  $i$ . Then, it can be shown that  $\varphi_d^{\text{KPCA}}(T_i, X_i) = \sqrt{n} V_{d,i}$ , and therefore,  $V^T \mathbf{e} / \sqrt{n} = \left( \hat{\mathbb{E}}_n [\varphi_1^{\text{KPCA}} \mathbf{e}], \dots, \hat{\mathbb{E}}_n [\varphi_n^{\text{KPCA}} \mathbf{e}] \right)^T$  and  $\tilde{V}^T \mathbf{e} / \sqrt{n} = \left( \hat{\mathbb{E}}_n [\varphi_1^{\text{KPCA}} \mathbf{e}], \dots, \hat{\mathbb{E}}_n [\varphi_D^{\text{KPCA}} \mathbf{e}] \right)^T$ . Using these relations, we can re-write the above quadratic forms as

$$\begin{aligned} \mathbf{e}^T M_{\text{GP}} \mathbf{e} &= \frac{1}{\sigma^2} \sum_{d=1}^n \frac{\lambda_d/n}{\lambda_d/n + \sigma^2/n} \left| \sqrt{n} \hat{\mathbb{E}}_n [\varphi_d^{\text{KPCA}}(T, X) \mathbf{e}(Y, T, X)] \right|^2 \\ \mathbf{e}^T M_{\text{low-rank GP}} \mathbf{e} &= \frac{1}{\sigma^2} \sum_{d=1}^D \frac{\lambda_d/n}{\lambda_d/n + \sigma^2/n} \left| \sqrt{n} \hat{\mathbb{E}}_n [\varphi_d^{\text{KPCA}}(T, X) \mathbf{e}(Y, T, X)] \right|^2. \end{aligned} \tag{24}$$

These expressions provide several intuitions of the kernel conditional moment constraints. First, the difference in the degree of freedom for the  $\chi^2$  upper bounds clearly corresponds to the number of summands for individual constraints. Second, when we decompose the individual summand into weight  $\frac{\lambda_d/n}{\lambda_d/n + \sigma^2/n}$  and squared penalty  $\left| \sqrt{n} \hat{\mathbb{E}}_n [\varphi_d^{\text{KPCA}}(T, X) \mathbf{e}(Y, T, X)] \right|^2$ , the weight term discounts the squared penalty as the magnitude of the spectrum  $\lambda_d$  decays. Especially, the summand becomes negligible when  $\lambda_d$  is significantly smaller than  $\sigma^2$ . If  $\lambda_D$  is small enough compared to  $\sigma^2$ , both quadratic forms are

approximately equal, and thus, the low-rank constraints can be tighter than its full-rank counterpart owing to the smaller  $\chi^2$  upper bound. This is a side benefit of low-rank constraints that can also justify its use in practice. Third, the squared penalty term can be interpreted as a soft version of constraint  $\mathbb{E}_{\text{obs}}[e(Y, T, X)\varphi_d^{\text{KPCA}}(T, X)] = 0$ . Lastly, we can approximately apply the central limit theorem to  $\sqrt{n}\hat{\mathbb{E}}_n[\varphi_d^{\text{KPCA}}(T, X)e(Y, T, X)]$  and obtain its asymptotic distribution. For the sake of approximation, let us assume that  $\mathbb{V}_{\text{obs}}[e(Y, T, X)|T, X] = \mathbb{E}_{\text{obs}}[|e(Y, T, X)|^2|T, X] = \sigma^2$  and that  $\varphi_d^{\text{KPCA}}(t, x)$  is fixed for any  $n$  and it satisfies  $\mathbb{E}_{\text{obs}}|\varphi_d^{\text{KPCA}}|^2 \approx 1$ . Then,  $\sqrt{n}\hat{\mathbb{E}}_n[\varphi_d^{\text{KPCA}}(T, X)e(Y, T, X)]$  is asymptotically normal with mean zero and variance  $\mathbb{V}_{\text{obs}}[\varphi_d^{\text{KPCA}}e] = \mathbb{E}_{\text{obs}}|\varphi_d^{\text{KPCA}}e|^2 = \mathbb{E}_{\text{obs}}[|\varphi_d^{\text{KPCA}}(T, X)|^2\mathbb{E}_{\text{obs}}[|e(Y, T, X)|^2|T, X]] = \sigma^2\mathbb{E}_{\text{obs}}|\varphi_d^{\text{KPCA}}(T, X)|^2 \approx \sigma^2$ . Therefore, we see that the above quadratic forms are the discounted sum of squares of asymptotically standard normal random variables, which again gives an intuition for the  $\chi^2$  upper bound.

### D.3 Choice of Parameter $\sigma^2 > 0$

Finally, we discuss a practice consideration on the choice of  $\sigma^2$ . From (24), we can see that the quadratic constraint gets tighter when we reduce the value of  $\sigma^2$ , as scaling factor  $\frac{1}{\sigma^2}$  of the summation, as well as discounting weight  $\frac{\lambda_d/n}{\lambda_d/n + \sigma^2/n}$ , gets larger as we decrease  $\sigma^2$ . In the sensitivity analysis, we are mostly interested in extreme cases, and thus, it is better to have overly loose constraints than overly tight constraints. Therefore, when choosing  $\sigma^2$ , we can take the largest possible value of  $\sup_{t \in \mathcal{T}, x \in \mathcal{X}} \mathbb{V}_{\text{obs}}[e|T = t, X = x]$ . In the case of the box constraints, we can calculate the upper and lower bound of  $e(y, t, x)$ , and therefore, the difference between both bounds can be a good guess of the supremum. With regard to the f-sensitivity model, there is no way to obtain such a bound. For this issue, a practical workaround may be to define the uncertainty set by combining both the f-divergence constraints and the box constraints, so that we can still use the choice of  $\sigma^2$  for the box constraints.

## E Theoretical Analysis

In this section, we study the property of the kernel conditional moment constraints in the confounding robust inference. For the convenience of the analysis, we only consider low-rank orthogonality condition  $\hat{\mathbb{E}}_n[e(Y, T, X)\psi(T, X)] = \mathbf{0}$  as the kernel conditional moment constraints in this section. Thus, we focus on the property of the following population lower bound

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

for

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

and its empirical version

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

for

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

We first study the property of the minimizers for the above problems. Then, we analyze the specification error of the kernel conditional moment constraints  $|V_{\text{inf}}^{\text{CMC}} - V_{\text{inf}}^{\text{KCMC}}|$ . We then provide a condition on orthogonal function set  $\{\psi_d(t, x)\}_{d=1}^D$  under which the specification error becomes zero. Finally, we study empirical estimator  $\hat{V}_{\text{inf}}^{\text{KCMC}}$  and prove consistency guarantees for both policy evaluation and learning.

Before further discussion, we introduce several simplifications of 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 parametrization  $\tilde{w}(y, t, x) = p_{\text{obs}}(t|x)w(y, t, x)$ .

Furthermore, we introduce 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  $f^*(v) := \sup_u \{uv - f(u)\}$ . There are a few important properties of the Fenchel



conjugate. The Fenchel conjugate is always convex because the supremum of a family of convex functions is convex.<sup>7</sup> Additionally, there exists maximizer  $u^*$  that solves  $f^*(v) = \max_u \{vu - f(u)\}$  and it satisfies

$$u^* \in \partial f^*(v) \quad (25)$$

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).

## E.1 Characterization of Solutions

In this section, 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

$$\begin{aligned} w_{\text{CMC}}^* &= \arg \min_{w \in \mathcal{W}^{\text{CMC}}} \mathbb{E}[w(Y, T, X)\pi(T|X)Y], \\ w_{\text{KCMC}}^* &= \arg \min_{w \in \mathcal{W}^{\text{KCMC}}} \mathbb{E}[w(Y, T, X)\pi(T|X)Y], \\ \hat{w}_{\text{KCMC}} &= \arg \min_{w \in \hat{\mathcal{W}}^{\text{KCMC}}} \hat{\mathbb{E}}_n[w(Y, T, X)\pi(T|X)Y]. \end{aligned} \quad (26)$$

Here, we know that these problems have the 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 (26). 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*

$$w_{\text{CMC}}^*(y, t, x) \in \left( \frac{1}{p_{\text{obs}}(t|x)} \right) \partial f_{t,x}^* \left( \frac{\eta_{\text{CMC}}(t, x) - r(y, t, x)}{\eta_f} \right), \quad (27)$$

$$w_{\text{KCMC}}^*(y, t, x) \in \left( \frac{1}{p_{\text{obs}}(t|x)} \right) \partial f_{t,x}^* \left( \frac{\eta_{\text{KCMC}}^T \psi(t, x) - r(y, t, x)}{\eta'_f} \right), \quad (28)$$

$$\hat{w}_{\text{KCMC}}(y, t, x) \in \left( \frac{1}{p_{\text{obs}}(t|x)} \right) \partial f_{t,x}^* \left( \frac{\eta'_{\text{KCMC}}^T \psi(t, x) - r(y, t, x)}{\eta''_f} \right). \quad (29)$$

**Proof** See below. □

### E.1.1 Characterization of $w_{\text{CMC}}^*$

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

$$\begin{aligned} V_{\text{inf}}^{\text{CMC}} &= \inf_{w \in \mathcal{W}^{\text{CMC}}} \mathbb{E}[\tilde{w}(Y, T, X)r(Y, T, X)] \\ &= \inf_{\tilde{w}} \sup_{\substack{\eta_{\text{CMC}}: \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}, \\ \eta_f \geq 0}} \mathbb{E}[\tilde{w}r] - \mathbb{E}[\eta_{\text{CMC}}(T, X)\mathbb{E}[\tilde{w} - 1|T, X]] + \eta_f (\mathbb{E}[f_{T,X}(\tilde{w})] - \gamma) \\ &= \sup_{\substack{\eta_{\text{CMC}}: \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}, \\ \eta_f > 0}} \inf_{\tilde{w}} \mathbb{E}[\tilde{w}r] - \mathbb{E}[\eta_{\text{CMC}}(T, X)(\tilde{w} - 1)] + \eta_f (\mathbb{E}[f_{T,X}(\tilde{w})] - \gamma) \\ &= \sup_{\substack{\eta_{\text{CMC}}: \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}, \\ \eta_f > 0}} -\eta_f \gamma + \mathbb{E}[\eta_{\text{CMC}}] - \eta_f \mathbb{E} \left[ \sup_{\tilde{w}} \left\{ \left( \frac{\eta_{\text{CMC}} - r}{\eta_f} \right) \tilde{w} - f_{T,X}(\tilde{w}) \right\} \right] \\ &= \sup_{\substack{\eta_{\text{CMC}}: \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}, \\ \eta_f > 0}} -\eta_f \gamma + \mathbb{E}[\eta_{\text{CMC}}] - \eta_f \mathbb{E} \left[ f_{T,X}^* \left( \frac{\eta_{\text{CMC}} - r}{\eta_f} \right) \right]. \end{aligned} \quad (30)$$

<sup>7</sup>In this case,  $v \mapsto vu - f(u)$  is linear, and therefore, is convex.

In the second last line, we assumed  $\eta_f > 0$ , as the inner minimization achieves  $-\infty$  if we take  $\eta_f = 0$  (unless  $r(Y, T, X)$  does not depend on  $Y$  almost surely).

Now, as primal solution  $w_{\text{CMC}}^*$  must satisfy the Karush–Kuhn–Tucker (KKT) conditions, we can take the maximizers of (30) as  $\eta_f^*$  and  $\eta_{\text{CMC}}^*(t, x)$ . Using property (25) 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 (27). Here, factor  $1/p_{\text{obs}}(t|x)$  in front of the subgradient appears because of reparametrization  $\tilde{w}(y, t, x) = p_{\text{obs}}(t|x)w(y, t, x)$ .

Now we are interested in dual solutions  $\eta_f^*$  and  $\eta_{\text{CMC}}^*$ . To characterize the dual solution, we take the stationary conditions for the dual problem as

$$\begin{aligned} 0 &\in \partial_{\eta_f} \left( -\eta_f \gamma + \mathbb{E}[\eta_{\text{CMC}}] - \eta_f \mathbb{E} \left[ f_{T,X}^* \left( \frac{\eta_{\text{CMC}} - r}{\eta_f} \right) \right] \right) \\ &= -\gamma - \mathbb{E} \left[ f_{T,X}^* \left( \frac{\eta_{\text{CMC}} - r}{\eta_f} \right) \right] + \mathbb{E} \left[ \left( \frac{\eta_{\text{CMC}} - r}{\eta_f} \right) \cdot \partial f_{T,X}^* \left( \frac{\eta_{\text{CMC}} - r}{\eta_f} \right) \right] \end{aligned}$$

and

$$\begin{aligned} 0 &\in \partial_{\eta_{\text{CMC}}} \left( -\eta_f \gamma + \mathbb{E}[\eta_{\text{CMC}}] - \eta_f \mathbb{E} \left[ f_{T,X}^* \left( \frac{\eta_{\text{CMC}} - r}{\eta_f} \right) \right] \right) \\ &= 1 - \mathbb{E} \left[ \partial f_{T,X}^* \left( \frac{\eta_{\text{CMC}} - r}{\eta_f} \right) \mid T = t, X = x \right], \end{aligned} \quad (31)$$

where we used the functional gradient on measure  $p_{\text{obs}}$  in the second subgradient condition. As  $\tilde{w}_{\text{CMC}}^*(y, t, x) := p_{\text{obs}}(t|x) \cdot w_{\text{CMC}}^*(y, t, x) \in \partial f_{t,x}^* \left( \frac{\eta_{\text{CMC}}^*(t,x) - r(y,t,x)}{\eta_f^*} \right)$ , we can see that the second condition corresponds to the conditional moment constraints.

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

**Example 2 (Solutions for box constraints)** *Let us consider the box constraints in the form 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, for this choice of function  $f_{t,x}$ , 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 subgradient term of condition (31), we can derive more explicit expression

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

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

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

where  $Q(t, x)$  was defined as the  $\tau(t, x)$ -th 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 (27) as

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

### E.1.2 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 above and reach a similar characterization of the solution as

$$\begin{aligned} V_{\text{inf}}^{\text{KCMC}} &= \inf_{w \in \mathcal{W}^{\text{KCMC}}} \mathbb{E}[\tilde{w}(Y, T, X)r(Y, T, X)] \\ &= \inf_{\tilde{w}} \sup_{\substack{\eta_{\text{KCMC}} \in \mathbb{R}^D, \\ \eta_f \geq 0}} \mathbb{E}[\tilde{w}r] - \mathbb{E}[(\tilde{w} - 1)\eta_{\text{KCMC}}^T \boldsymbol{\psi}] + \eta_f (\mathbb{E}[f_{T,X}(\tilde{w})] - \gamma) \\ &= \sup_{\substack{\eta_{\text{KCMC}} \in \mathbb{R}^D, \\ \eta_f \geq 0}} \mathbb{E}[\tilde{w}r] - \mathbb{E}[(\tilde{w} - 1)\eta_{\text{KCMC}}^T \boldsymbol{\psi}] + \eta_f (\mathbb{E}[f_{T,X}(\tilde{w})] - \gamma) \\ &= \sup_{\substack{\eta_{\text{KCMC}} \in \mathbb{R}^D, \\ \eta_f > 0}} -\eta_f \gamma + \eta_{\text{KCMC}}^T \mathbb{E}[\boldsymbol{\psi}] - \eta_f \mathbb{E} \left[ \sup_{\tilde{w}} \left\{ \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r}{\eta_f} \right) \tilde{w} - f_{T,X}(\tilde{w}) \right\} \right] \\ &= \sup_{\substack{\eta_{\text{KCMC}} \in \mathbb{R}^D, \\ \eta_f > 0}} -\eta_f \gamma + \eta_{\text{KCMC}}^T \mathbb{E}[\boldsymbol{\psi}] - \eta_f \mathbb{E} \left[ f_{T,X}^* \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r}{\eta_f} \right) \right]. \end{aligned} \quad (32)$$

Now, using the maximizers of dual problem (32)  $\eta_f^*$  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 (28).

Now we study the characterization of dual solutions  $\eta_f^*$  and  $\eta_{\text{KCMC}}^*$ . Again, we take the stationary conditions as

$$\begin{aligned} 0 &\in \partial_{\eta_f} \left( -\eta_f \gamma + \eta_{\text{KCMC}}^T \mathbb{E}[\boldsymbol{\psi}] - \eta_f \mathbb{E} \left[ f_{T,X}^* \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r}{\eta_f} \right) \right] \right) \\ &= -\gamma - \mathbb{E} \left[ f_{T,X}^* \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r}{\eta_f} \right) \right] + \mathbb{E} \left[ \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r}{\eta_f} \right) \cdot \partial f_{T,X}^* \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r}{\eta_f} \right) \right] \end{aligned}$$

and

$$\begin{aligned} 0 &\in \partial_{\eta_{\text{KCMC}}} \left( -\eta_f \gamma + \eta_{\text{KCMC}}^T \mathbb{E}[\boldsymbol{\psi}] - \eta_f \mathbb{E} \left[ f_{T,X}^* \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r}{\eta_f} \right) \right] \right) \\ &= \mathbb{E} \left[ \boldsymbol{\psi} \left( 1 - \partial f_{T,X}^* \left( \frac{\eta_{\text{KCMC}}^T \boldsymbol{\psi} - r}{\eta_f} \right) \right) \mid T = t, X = x \right]. \end{aligned}$$

Again, as  $\tilde{w}_{\text{KCMC}}^*(y, t, x) = p_{\text{obs}}(t|x) \cdot w_{\text{KCMC}}^*(y, t, x) \in \partial f_{t,x}^* \left( \frac{\eta_{\text{KCMC}}^{*T} \boldsymbol{\psi}(t,x) - r(y,t,x)}{\eta_f^*} \right)$ , we can see that the second condition corresponds to the kernel conditional moment constraints.

### E.1.3 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

$$\sup_{\substack{\eta_{\text{KCMC}} \in \mathbb{R}^D, \\ \eta_f > 0}} -\eta_f \gamma + \eta_{\text{KCMC}}^T \hat{\mathbb{E}}_n[\psi] - \eta_f \hat{\mathbb{E}}_n \left[ f_{T,X}^* \left( \frac{\eta_{\text{KCMC}}^T \psi - r}{\eta_f} \right) \right]. \quad (33)$$

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

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

which proves (29).

## E.2 Specification Error

Using the above characterization of the solutions, we can find a condition under which the specification error of estimator  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi)$  becomes zero for policy  $\pi$  so that  $|V_{\text{inf}}^{\text{KCMC}}(\pi) - V_{\text{inf}}^{\text{CMC}}(\pi)| = 0$ .

**Theorem 1 (No specification error)** *Let  $\eta_{\text{CMC}}^*(t, x)$  be the solution of dual problem (30) for  $V_{\text{inf}}^{\text{CMC}}(\pi)$ . Then, if*

$$\eta_{\text{CMC}}^* \in \text{span}(\{\psi_1, \dots, \psi_D\}), \quad (34)$$

we have

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

**Proof** Take  $\eta_f^*$  such that  $(\eta_{\text{CMC}}^*(t, x), \eta_f^*)$  is the solution of dual problem (30) for  $V_{\text{inf}}^{\text{CMC}}$ . Then, we can take multiplier  $\eta_{\text{KCMC}}^*$  that satisfies  $\eta_{\text{CMC}}^* = \eta_{\text{KCMC}}^* \psi$ . Now, we can see that dual problem (32) for  $V_{\text{inf}}^{\text{KCMC}}$  can be considered as the restricted version of dual problem (30) 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 (32). 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  $V_{\text{inf}}^{\text{CMC}} = V_{\text{inf}}^{\text{KCMC}}$ .  $\square$

Interestingly, with the above result, we can derive quantile balancing constraint (23) for the previously proposed sharp estimator by Dorn and Guo (2022).

**Example 3 (Derivation of quantile balancing estimator (Dorn and Guo, 2022))** *Let us consider the same box constraints as Example 2. For this problem, we know the analytical form of dual solution  $\eta_{\text{CMC}}^*(t, x) = \left( \frac{\pi(t|x)}{p_{\text{obs}}(t|x)} \right) Q(t, x)$ .*

*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 (34) in Theorem 1 to obtain the kernel conditional moment constraint with no specification error.*

## E.3 Consistency of Policy Evaluation and Learning

Lastly, 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 reduction of our problem to the M-estimation (Van de Geer, 2000) using the dual formulation. We similarly show a reduction of policy learning problem  $\max_{\pi \in \Pi} \hat{V}_{\text{inf}}^{\text{KCMC}}$  to the M-estimation and prove the consistency of the learned policy parameter when policy class is finite-dimensional and concave. This reduction to the M-estimation significantly simplified our proof compared to the proof in Kallus and Zhou (2021) using uniform convergence, because we can take advantage of the well-studied theory of M-estimation. Though their approach using uniform convergence is very powerful, we found it not immediately applicable to our work, due to the difficulty of taking the empirical moment constraints into account.

### E.3.1 Consistency of Policy Evaluation

To prove the consistency of policy evaluation and learning, we will make use of the two following convergence lemmas for loss function  $\ell : \Theta \times \mathcal{Z} \rightarrow \mathbb{R}$ , where  $\Theta \subseteq \mathbb{R}^K$  for some  $K$  and  $\mathcal{Z} := \mathcal{Y} \times \mathcal{T} \times \mathcal{X}$ . We assume that for  $Z := (Y, T, X)$ , the loss function satisfies  $\mathbb{E}|\ell_\theta(Z)| < \infty$  for any  $\theta \in \Theta$ .

**Lemma 2 (Uniform convergence on compact space, Van de Geer (2000, Lemma 5.2.2.))** Assume that parameter space  $(\Theta, \|\cdot\|)$  is compact. Also assume that  $\theta \mapsto \ell_\theta$ ,  $\theta \in \Theta$  is continuous and it has an  $L^1$  envelope so that  $\mathbb{E}[G(Z)] < \infty$  for  $G(z) := \sup_{\theta \in \Theta} |\ell_\theta(z)|$ . Then,  $\sup_{\theta \in \Theta} \left| \hat{\mathbb{E}}_n[\ell_\theta(Z)] - \mathbb{E}[\ell_\theta(Z)] \right| \xrightarrow{P} 0$ .

**Lemma 3 (Consistency of convex M-estimation, Van de Geer (2000, Lemma 5.2.3.))** Let us define  $\theta_0 \in \arg \min_{\theta \in \Theta} \mathbb{E}[\ell_\theta(Z)]$  and its M-estimator  $\hat{\theta}_n \in \arg \min_{\theta \in \Theta} \hat{\mathbb{E}}_n[\ell_\theta(Z)]$ . Suppose that  $\theta_0$  is the unique minimizer and that  $\Theta \subseteq \mathbb{R}^k$  is convex. Also assume that  $\theta \mapsto \ell_\theta$ ,  $\theta \in \Theta$  is continuous and convex, satisfying  $\mathbb{E}[G_\varepsilon] < \infty$  for  $G_\varepsilon(z) := \sup_{\theta \in \Theta: \|\theta - \theta_0\| \leq \varepsilon} |\ell_\theta(z)|$  for some  $\varepsilon > 0$ . Then,  $\hat{\theta}_n \xrightarrow{P} \theta_0$ .

As our dual problem for policy evaluation (32) and (33) are concave maximization, we can immediately apply the above lemma as follows.

**Theorem 2 (Consistency of policy evaluation)** Define parameter space of  $(\eta_f, \eta_{\text{KCMC}})$  as  $\Theta \subseteq \mathbb{R}_+ \times \mathbb{R}^D$ . Further, define  $\theta_0 := (\eta_f^*, \eta_{\text{KCMC}}^*)$  as the solution of dual problem (32) for  $V_{\text{inf}}^{\text{KCMC}}$  and  $\hat{\theta}_n := (\hat{\eta}_f, \hat{\eta}_{\text{KCMC}})$  as the solution to dual problem (33) for  $\hat{V}_{\text{inf}}^{\text{KCMC}}$ . Now, assume that  $\theta_0$  is unique and that  $\theta_0 \in \Theta$ . Also assume that  $f_{t,x}^* : \mathbb{R} \rightarrow \mathbb{R}$  is continuous for any  $t \in \mathcal{T}$  and  $x \in \mathcal{X}$ . 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 the inside of the expectation of dual objectives (32) and (33). Furthermore, assume  $\ell$  satisfies  $\mathbb{E}|\ell_\theta| < \infty$  for any  $\theta \in \Theta$  and  $\mathbb{E}[G_\varepsilon] < \infty$  for  $G_\varepsilon := \sup_{\theta \in \Theta: \|\theta - \theta_0\| \leq \varepsilon} |\ell_\theta|$  for some  $\varepsilon > 0$ . Then, we have  $\hat{\theta}_n \xrightarrow{P} \theta_0$  and  $\hat{V}_{\text{inf}}^{\text{KCMC}} \xrightarrow{P} V_{\text{inf}}^{\text{KCMC}}$ .

**Proof** We can immediately apply Lemma 3 and get  $\hat{\theta}_n \xrightarrow{P} \theta_0$ . Thus,  $\hat{\theta}_n$  tend to the inside of compact set  $\{\theta \in \Theta : \|\theta - \theta_0\| \leq \varepsilon\}$ , in which we have the uniform convergence of  $\hat{\mathbb{E}}_n[\ell_\theta(Z)]$  to  $\mathbb{E}[\ell_\theta(Z)]$ . Therefore, we have  $\hat{V}_{\text{inf}}^{\text{KCMC}} = -\hat{\mathbb{E}}_n[\ell_{\hat{\theta}_n}(Z)] \xrightarrow{P} -\mathbb{E}[\ell_{\hat{\theta}_n}(Z)] \xrightarrow{P} -\mathbb{E}[\ell_{\theta_0}(Z)] = V_{\text{inf}}^{\text{KCMC}}$ .  $\square$

In practice, it is difficult to check the assumption of integrability condition  $\mathbb{E}|\ell_\theta| < \infty$  for any  $\theta \in \Theta$  as well as the uniqueness of the solution. However, it is possible in some cases to check  $L^1$  envelope condition  $\mathbb{E}[G_\varepsilon] < \infty$ , because local Lipschitzness of  $f^*$  implies the existence of such  $\varepsilon$ . For example, for the box constraints of Example 2, we know that  $f^*$  is upper bounded by  $b_{\bar{w}}(t, x)$ . For the f-divergence constraints, the conjugate function  $f^*$  for many choices of f-divergence is locally Lipschitz, as shown in Table 1.

### E.3.2 Consistency of Concave Policy Learning by M-estimation

Now we consider policy learning. Instead of providing the standard uniform convergence proof, our theoretical result leverages the preceding lemmas.

**Theorem 3 (Consistency of concave policy learning)** 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  $x \in \mathcal{X}$ . Define the parameter space of  $\theta = (\beta, \eta_f, \eta_{\text{KCMC}})$  as  $\Theta = \mathcal{B} \times \Theta_\eta$  for some convex  $\Theta_\eta \subseteq \mathbb{R}_+ \times \mathbb{R}^D$ . Define also  $\theta_0 := (\beta^*, \eta_f^*, \eta_{\text{KCMC}}^*)$  so that  $\beta^* \in \arg \max_{\beta \in \mathcal{B}} V_{\text{inf}}^{\text{KCMC}}(\pi_\beta)$  and  $(\eta_f^*, \eta_{\text{KCMC}}^*)$  is the solution of dual problem (32) for  $V_{\text{inf}}^{\text{KCMC}}$  at policy  $\pi_{\beta^*}$ . Similarly, define  $\hat{\theta}_n := (\hat{\beta}, \hat{\eta}_f, \hat{\eta}_{\text{KCMC}})$  so that  $\hat{\beta} \in \arg \max_{\beta \in \mathcal{B}} \hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_\beta)$  and  $(\hat{\eta}_f, \hat{\eta}_{\text{KCMC}})$  is the solution of dual problem (33) for  $\hat{V}_{\text{inf}}^{\text{KCMC}}$  at policy  $\pi_{\hat{\beta}}$ . Now, assume that  $\theta_0$  is unique and that  $\theta_0 \in \Theta$ . Also assume that  $f_{t,x}^* : \mathbb{R} \rightarrow \mathbb{R}$  and  $\beta \rightarrow \pi_\beta$  is continuous for any  $t \in \mathcal{T}$  and  $x \in \mathcal{X}$ . Define  $\ell : \Theta \times \mathcal{Z} \rightarrow \mathbb{R}$  as

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

so that it is the negative version of the inside of the expectation of dual objectives (32) and (33). Furthermore, assume  $\ell$  satisfies  $\mathbb{E}|\ell_\theta| < \infty$  for any  $\theta \in \Theta$  and  $\mathbb{E}[G_\varepsilon] < \infty$  for  $G_\varepsilon := \sup_{\theta \in \Theta: \|\theta - \theta_0\| \leq \varepsilon} |\ell_\theta|$  for some  $\varepsilon > 0$ . Then, we have  $\hat{\theta}_n \xrightarrow{P} \theta_0$  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

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

is a concave maximization problem, because  $\mathbb{E}[\ell_{\theta}]$  is the pointwise infimum of concave functions. Thus, we can apply Lemma 3 and get  $\hat{\theta}_n \xrightarrow{P} \theta_0$ , which implies  $\hat{\theta}_n$  tend to the inside of compact set  $\{\theta \in \Theta : \|\theta - \theta_0\| \leq \varepsilon\}$ , where we have uniform convergence guarantee of  $\hat{\mathbb{E}}_n[\ell_{\theta}(Z)]$ . Therefore, we have  $\hat{V}_{\text{inf}}^{\text{KCMC}}(\pi_{\hat{\beta}}) = -\hat{\mathbb{E}}_n[\ell_{\hat{\theta}_n}(Z)] \xrightarrow{P} -\mathbb{E}[\ell_{\hat{\theta}_n}(Z)] \xrightarrow{P} -\mathbb{E}[\ell_{\theta_0}(Z)] = V_{\text{inf}}^{\text{KCMC}}(\pi_{\beta^*})$ .  $\square$

## F Experimental Settings and Additional Numerical Examples

Lastly, we provide the details of our numerical experiments and provide more experimental results of the f-sensitivity models.

### F.1 Datasets

In the experiment, we used two types of data, one is synthetic and the other is real-world data. We base most of the experiments on the first synthetic data adopted from Kallus and Zhou (2018, 2021). The second dataset is a real-world example used in Dorn and Guo (2022), and it is used to illustrate the application of our methods to a real-world dataset.

Now we explain our first dataset. We use the following data-generating process for this synthetic data.

$$\begin{aligned}
 \xi &\sim \text{Bern}(1/2), \\
 X &\sim \mathcal{N}(\mu_x, \mathbf{I}_5), \\
 Y_0|X, \xi &\sim \mathcal{N}(\beta_{x,0}^T X + \beta_{\xi,0}\xi + \beta_{\text{const},0}, 1), \\
 Y_1|X, \xi &\sim \mathcal{N}(\beta_{x,1}^T X + \beta_{\xi,1}\xi + \beta_{\text{const},1}, 1), \\
 U &= \mathbf{1}_{Y_0 > Y_1}, \\
 T|X, Y_1, Y_2, U &\sim \text{Bern}(e(X, U)), \\
 Y &= Y_T,
 \end{aligned}$$

where  $e(X, U) := \frac{6e(X)}{4+5U+e(X)(2-5U)}$  and  $e(x) := \sigma(\beta_e^T x)$ . Here,  $\sigma(u) := \frac{\exp(u)}{1+\exp(u)}$  indicates the sigmoid function. The parameters we used are

$$\begin{aligned}
 \mu_x &= [-1, 0.5, -1, 0, -1], \\
 \beta_{x,0} &= [0, .5, -0.5, 0, 0], \\
 \beta_{x,1} &= [-1.5, 1.5, -2, 1, 0.5], \\
 \beta_{\xi,0} &= 1, \\
 \beta_{\xi,1} &= -1, \\
 \beta_{\text{const},0} &= 2.5, \\
 \beta_{\text{const},1} &= -0.5, \\
 \beta_e &= [0, 0.75, -0.5, 0, -1].
 \end{aligned}$$

For the policy evaluation task, we used policy  $\pi(t = 1|x) := e(x)$ .

For the real-world data example, we use the same dataset as Dorn and Guo (2022), which is 668 subsamples of data from the 1966-1981 National Longitudinal Survey (NLS) of Older and Young Men. The subsamples consist of the 1978 cross-section

of Young Men who are craftsmen or laborers and are not enrolled in school. We estimate the average treatment effect ( $\mathbb{E}[Y|T = 1] - \mathbb{E}[Y|T = 0]$ ) of union membership ( $T$ ) on log wages ( $Y$ ), and eight other covariates are used as  $X$ . For the average treatment effect estimation, we substitute  $\mathbb{1}_{t=1} - \mathbb{1}_{t=0}$  in place of  $\pi(t|x)$ . Note that this substituted quantity is a difference of policy  $\pi_1(t|x) := \mathbb{1}_{t=1}$  and  $\pi_0(t|x) := \mathbb{1}_{t=0}$ , and therefore it is not a proper policy. Nevertheless, such a substitution is still possible, as our method can accommodate any function  $\pi(t|x)$  in place of the evaluated policy.<sup>8</sup>

For synthetic data, we generate a dataset of 500 samples for individual experimental configurations, unless otherwise specified. We repeat the experiment 10 times using different random seeds and report the mean of the 10 experiments. Additionally, we indicate plus/minus one standard deviation from the mean by the colored band around the line representing the mean value.

Lastly, conditional probability  $p_{\text{obs}}(t|x)$  used to construct the estimators was estimated from the data using the logistic regression with covariate  $X$ .

## F.2 Compared Estimators

In the numerical examples, we consider four types of estimators.

First, as the baseline method, we consider the conventionally used ZSB estimator. We impose the ZSB constraints on other estimators, in order to see the additional improvements by these constraints. To impose ZSB constraints while ensuring the feasibility of the associated convex programming, we applied appropriate rescaling to the estimates of  $p_{\text{obs}}(t|x)$ .<sup>9</sup> As discussed above, this estimator is conventionally used (Tan, 2006; Kallus and Zhou, 2018; Zhao et al., 2019) but is known to provide too conservative bounds.

Against this baseline, we compared two types of the proposed estimators based on the kernel moment constraints (KCMC), which are the low-rank Gaussian process KCMC and the low-rank hard KCMC using the orthogonal function class obtained by the kernel PCA. In the following, we call them "low-rank GP KCMC" and "low-rank hard KCMC", respectively. For both of the low-rank KCMC estimators, we used 100-rank approximation so that  $D = 100$ , unless otherwise specified.

Additionally, we compared the quantile balancing (QB) estimator by Dorn and Guo (2022). As discussed in Example 3, this estimator can be considered as a special case of the low-rank hard KCMC estimators that uses the optimal orthogonal function class, in the case of box constraints.

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).

## F.3 Additional Numerical Experiments with f-divergence Sensitivity Models

Here, we list more examples of the f-sensitivity analysis with various types of f-divergences using the synthetic dataset. The f-divergences considered here are listed in Table 1.

Table 1: Commonly used f-divergence, corresponding convex function  $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ , and its Fenchel conjugate  $f^* : \text{dom}(f^*) \rightarrow \mathbb{R}$ .

| f-divergence      | $f(u)$  | $f^*(v)$             | $\text{dom}(f^*)$                      |
|-------------------|---|----------------------|--|
| KL                | $u \log u$  | $\exp(v - 1)$        | $\mathbb{R}$                           |
| Reverse KL        | $-\log u$   | $-1 - \log(-v)$      | $\mathbb{R}_-$                         |
| Jensen-Shannon    | $-(u + 1) \log \left(\frac{u+1}{2}\right) + u \log u$ | $-\log(2 - \exp(v))$ | $v < \log 2$                           |
| Squared Hellinger | $(\sqrt{u} - 1)^2$                                    | $\frac{v}{1-v}$      | $v < 1$                                |
| Pearson $\chi^2$  | $(u - 1)^2$   | $\frac{1}{4}v^2 + v$ | $\mathbb{R}$                           |
| Neyman $\chi^2$   | $\frac{1}{u} - 1$                                     | $-2\sqrt{-v} + 1$    | $\mathbb{R}_-$                         |
| Total Variation   | $\frac{1}{2} u - 1 $                                  | $v$                  | $-\frac{1}{2} \leq v \leq \frac{1}{2}$ |

<sup>8</sup>If we re-define the reward and the evaluated policy as  $Y' := 2(\mathbb{1}_{T=1}Y - \mathbb{1}_{T=0}Y)$  and  $\pi(t|x) := \frac{1}{2}$ , the offline policy evaluation of such a setup is equivalent to the average treatment effect estimation.

<sup>9</sup>We multiplied estimate  $\hat{p}_{\text{obs}}(t|X_i)$  by  $\frac{1}{n} \sum_{i=1}^n [\mathbb{1}_{T_i=t} / \hat{p}_{\text{obs}}(T_i|X_i)]$ . When all the constraints can be expressed as linear constraints, fractional linear programming can be used to impose the ZSB constraints in a more natural manner (Zhao et al., 2019). However, the quadratic constraints of GP KCMC make it impossible to use the linear fractional programming approach.

Similarly to the case of Tan’s marginal sensitivity models, the sharp estimators are tighter than the ZSB estimator. The low-rank hard KCMC is also providing (potentially excessively) tighter bounds than the low-rank Gaussian process KCMC and the quantile balancing estimators.

Interestingly, the quantile balancing constraint <sup>10</sup> is providing almost as sharp bound as the GP KCMC-based methods, even though it is no longer the theoretically optimal constraint. Still, it is possible to give some intuition on the use of orthogonal function class  $\{\varphi_1\}$  with  $\varphi_1(t, x) := \left(\frac{\pi(t|x)}{p_{\text{obs}}(t|x)}\right) \hat{Y}(t, x)$  as follows: If we have a good regressor of  $Y$  satisfying  $\hat{Y}(T, X) \approx Y$  and if  $\tilde{w}$  follows constraint  $\mathbb{E}_{\text{obs}} \left[ (\tilde{w}(Y, T, X) - 1) \left(\frac{\pi(T|X)}{p_{\text{obs}}(T|X)}\right) \hat{Y}(T, X) \right] = 0$ , we have

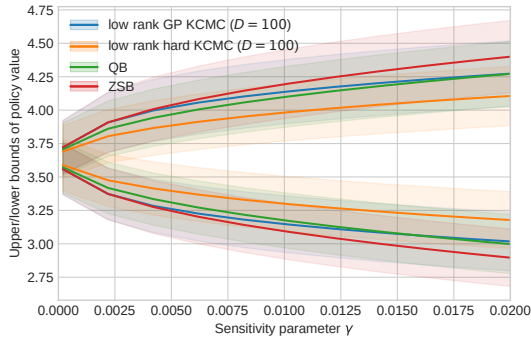
$$\begin{aligned} & \mathbb{E}_{\text{obs}} \left[ \tilde{w}(Y, T, X) \left(\frac{\pi(T|X)}{p_{\text{obs}}(T|X)}\right) Y(T, X) \right] \\ &= \mathbb{E}_{\text{obs}} \left[ \tilde{w}(Y, T, X) \left(\frac{\pi(T|X)}{p_{\text{obs}}(T|X)}\right) \{(Y - \hat{Y}(T, X)) + \hat{Y}(T, X)\} \right] \\ &= \mathbb{E}_{\text{obs}} \left[ \tilde{w}(Y, T, X) \left(\frac{\pi(T|X)}{p_{\text{obs}}(T|X)}\right) (Y - \hat{Y}(T, X)) \right] + \mathbb{E}_{\text{obs}} \left[ \left(\frac{\pi(T|X)}{p_{\text{obs}}(T|X)}\right) \hat{Y}(T, X) \right] \\ &\approx \mathbb{E}_{\text{obs}} \left[ \left(\frac{\pi(T|X)}{p_{\text{obs}}(T|X)}\right) \hat{Y}(T, X) \right]. \end{aligned}$$

This implies that if  $Y$  is easy to predict with  $X$  and  $T$ , the orthogonal constraint using the regressor  $\hat{Y}(t, x)$  gives a similar value to the IPW estimator of  $\hat{Y}(T, X)$  as long as  $\tilde{w}$  is not too far from 1.

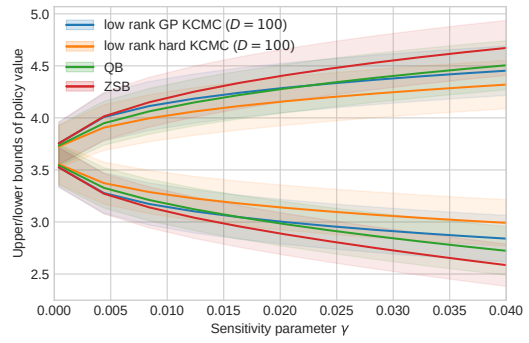
---

<sup>10</sup>We used  $\Gamma = 1.5$ , which corresponds to finding  $\hat{Y}(t, x)$  that approximates the 40 percentile.

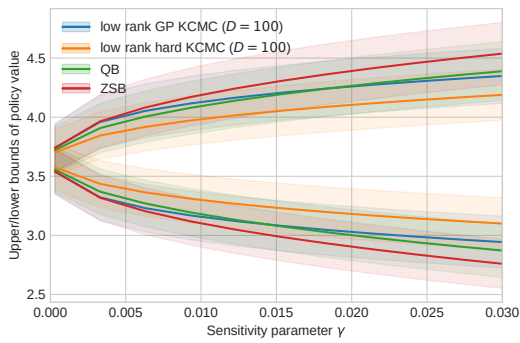




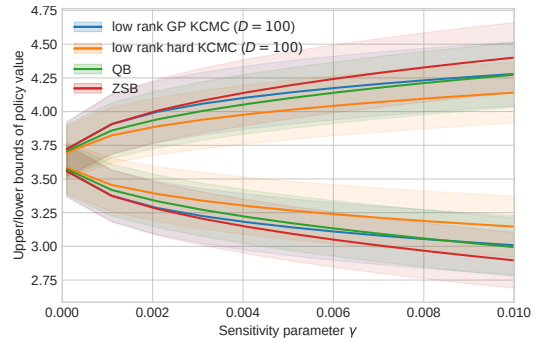
(a) KL sensitivity model



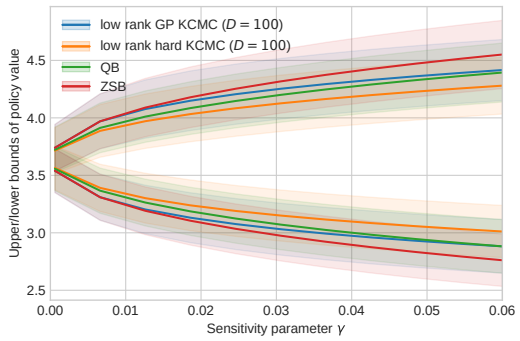
(b) Reverse KL sensitivity model



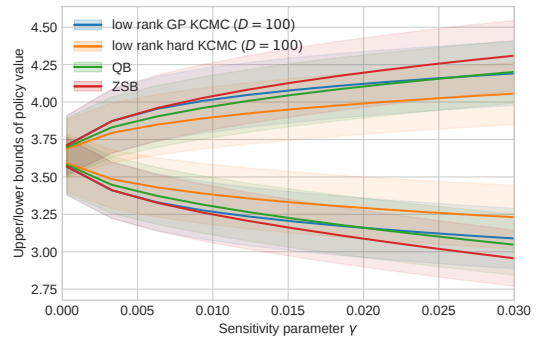
(c) Jensen-Shannon sensitivity model



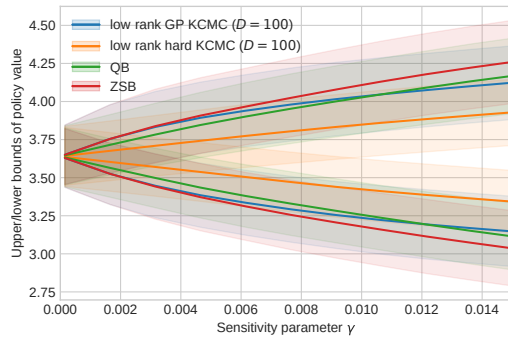
(d) Squared Hellinger sensitivity model



(e) Pearson  $\chi^2$  sensitivity model



(f) Neyman  $\chi^2$  sensitivity model



(g) Total variation sensitivity model

Figure 3: Estimated upper and lower bounds of policy value using different estimators for different f-sensitivity models. The synthetic data is used.