Finite Dimensional Approximations of Incentive Functions for Causal Incentive Design Models

Sebastián Bejos[1] SEBASTIAN.BEJOS@INAOEP.MX
Eduardo F. Morales[1] EMORALES@INAOEP.MX
Enrique Munoz de Cote[2] ENRIQUE@MUTABLETACTICS.AI
Luis Enrique Sucar[1] ESUCAR@INAOEP.MX
[1] *Computer Science Department, Instituto Nacional de Astrofísica Óptica y Electrónica; Tonantzintla, Puebla, México.*

[2] *Mutable Tactics; London CB2 9PJ, UK.*

Editors: Gustau Camps-Valls, Manuele Leonelli and Gherardo Varando

Abstract

Single-stage principal-agent problems with bilateral private information can be represented as additive-noise Causal Incentive Design Models (CIDMs), where an incentive policy γ acts as a functional intervention on a function-valued variable Γ. This research extends the Causal Incentive Design framework to address the computational challenge of densely sampling CIDMs when the intervention space Γ is a Reproducing Kernel Hilbert Space (RKHS). This challenge arises because RKHS feature maps are typically infinite-dimensional, while CIDM simulation requires that each incentive policy be instantiated as an explicit callable function and evaluated within a structural equation involving the follower's best-response optimization. To overcome this, quasi-Monte Carlo (QMC) and MC Random Fourier Features (RFF) frameworks are used to construct finite-dimensional approximations of RKHS-valued incentive functions and to densely sample CIDMs. The developed samplers are evaluated on five distinct CIDM instantiations. The experiments compare MC-RFF and QMC-RFF in terms of the approximation error of the RKHS kernel and the subsequent estimation of the causal policy-value target $V(\gamma)$ using Functional Kernel Ridge Regression and a Functional Multi-Layer Perceptron, benchmarked against an analytical Gaussian-smoothed baseline.

Keywords: Causal Inference, Statistical Contract Theory, Principal-agent Models, Kernel Methods, Sampling Function Spaces, Incentive Design, Functional Data Analysis.

1. Introduction

Incentive design within principal–agent systems requires reasoning about how a principal's selection of a contract, reward, or policy function influences the behavior of a strategic follower. This problem becomes especially challenging under bilateral private information, where each party possesses private knowledge of its own utility function and the principal lacks direct access to the follower's best-response mechanism. In the single-stage principal–agent setting considered in this work, the principal selects an incentive function γ, the follower responds by choosing an action that maximizes its private utility, and the principal observes only historical interaction tuples of incentive policies, follower actions, principal actions, and realized utilities. Causal Incentive Design Models (CIDMs) offer a formal causal graphical framework for analyzing such systems by conceptualizing incentive policies as interventions in a function-space variable Γ.

© 2026 S. Bejos, E.F. Morales, E. Munoz de Cote & L.E. Sucar.

Traditional inference and simulation frameworks for Causal Graphical Model (CGM) are primarily designed for discrete or finite-dimensional continuous variables. However, CIDMs present a non-standard computational challenge: the intervention space Γ may be infinite-dimensional when represented as a Reproducing Kernel Hilbert Space (RKHS). Simulating a CIDM necessitates that incentive policies be instantiated as explicit callable functions, yet exact RKHS feature maps cannot be directly sampled or stored in finite memory.

From the CGMs perspective, the principal difficulty extends beyond the size of the policy space Γ. The root variable Γ in a CIDM takes values in a function space, and each realized policy must be evaluated within two downstream structural mechanisms, including the nested optimization that determines the follower's best response. As a result, standard ancestral simulation requires an explicit and callable realization of the entire intervention function, rather than a scalar or finite-dimensional node value.

This representation problem is addressed by introducing finite-dimensional approximations of incentive functions using Random Fourier Features (RFF), enabling policies sampled from an RKHS-like function class to be represented as tractable linear combinations of sinusoidal features. The downstream Causal Graphical Model (CGM) and its conditional independence structure remain unchanged; only the representation of the function-valued root variable Γ and the interventions $do(\Gamma = \gamma)$ are approximated. Incorporating Quasi-Monte Carlo Random Fourier Features into the generative sampling process provides a mathematically rigorous and computationally tractable framework for managing these functional interventions. This methodology extends the modeling and inference capabilities of CGMs to complex, strategic policy spaces while mitigating the curse of dimensionality.

This work makes three primary contributions. First, it formalizes the finite-feature CIDM induced by a random Fourier feature (RFF) representation of the incentive policy. Second, Monte Carlo and randomized quasi-Monte Carlo spectral constructions are developed to generate callable policies and propagate them through the CIDM structural equations. Third, it evaluates the approximations at both the kernel-representation level and the downstream policy-value estimation level.

2. Single-Stage Causal Incentive Design Models

The standard semantics of an acyclic structural causal model $\mathcal{M} = \langle \mathbf{U}, \mathbf{V}, \mathbf{F}, p(\mathbf{U}) \rangle$ with mutually independent exogenous variables are employed. An intervention $do(\Gamma = \gamma)$ sets the deterministic policy root Γ to the incentive function γ. Generic definitions for structural causal models, Markov factorization, and interventions follow Pearl (2009). The following presents only the mechanisms and interventional quantities specific to Causal Incentive Design Models.

2.1. Principal-agent Models

Let $\omega_p \in \mathbf{\Omega}_p \subseteq \mathbb{R}^{n_p}$, with $n_p \in \mathbb{N}$ and $p \in \{L, F\}$, be the decision variables of the leader and follower and $J_p : \mathbf{\Omega}_L \times \mathbf{\Omega}_F \to \mathbb{R}$ their respective utility functions. The inverse Stackelberg problem involves the leader initially identifying a specific desired equilibrium $\left(\omega_L^d, \omega_F^d\right)$ to achieve; such as the global optimum of the leader's utility function:

$$\big(\omega_L^d, \omega_F^d\big) \in \underset{(\omega_L, \omega_F) \in \Omega_L \times \Omega_F}{\arg\max} J_L\big(\gamma(\omega_F), \omega_F\big),$$

for some incentive function $\gamma \in \Gamma$ in a function space $\Gamma : \mathbf{\Omega}_F \to \mathbf{\Omega}_L$. Given this target equilibrium, the remaining problem can be formulated as the following functional bi-level functional optimization problem (see Groot et al. (2013, 2017)):

$$\text{Find:} \qquad \gamma^* \in \underset{\gamma \in \Gamma}{\arg\max}\, J_L(\gamma(\omega_F), \omega_F) \tag{1}$$

$$\text{s.t.} \qquad (\omega_L^d, \omega_F^d) \in \underset{\omega_F \in \mathbf{\Omega}_F}{\arg\max}\, J_F(\gamma(\omega_F), \omega_F), \tag{2}$$

$$\gamma(\omega_F^d) = \omega_L^d. \tag{3}$$

A Single-Stage Principal-Agent Problem (SS-PAP) can be formulated as an inverse Stackelberg game, with an explicit information model that specifies the extent of knowledge each party has regarding the other's utility function. This study examines SS-PAPs under a bilateral private information model, where both the principal and the follower retain private information about their respective utility functions. In particular, the follower's utility function remains unknown to the principal, and the principal's utility function is similarly unknown to the follower. Within this framework, the principal selects an incentive function γ from a predetermined set $\Gamma = \{\gamma \mid \gamma : \mathbf{\Omega}_F \to \mathbf{\Omega}_L\}$ and communicates γ to the follower, without access to the follower's utility function. The follower, upon receiving the incentive function γ, then selects an action $\omega_F{}^* \in \mathbf{\Omega}_F$ that maximizes its utility $J_F(\gamma(\omega_F), \omega_F)$. The outcome of the single-stage principal-agent problem is characterized by the utilities $J_F(\omega_F^*), \omega_F^*)$ and $J_L(\omega_F^*), \omega_F^*)$ following the principal's choice of the incentive function γ.

Our research investigates the principal's decision-making process in selecting an incentive function γ that maximizes its utility, given access to historical data comprising recorded tuples $(\gamma^i, \omega_F^i, \omega_L^i, j_L^i)$ for $i \in [N]$ with $N \in \mathbb{N}$. That is, it is assumed that a data set $\mathcal{D} = \big\{\gamma^i, \omega_F^i, \omega_L^i, j_L^i\big\}_{i=1}^N$ is available. The primary focus of our research is on integrating causal inference methods into this class of problems. `Figure 2` in Appendix D presents a schematic abstraction of SS-PAPs when historical data is used as input.

2.2. Causal Incentive Design Models

Let $PAP_{L_1F_1}$ be the canonical SS-PAP involving one principal and one follower, each controlling a single variable, so $\omega_L \in \mathbf{\Omega}_L \subseteq \mathbb{R}$ and $\omega_F \in \mathbf{\Omega}_F \subseteq \mathbb{R}$, where $\mathbf{\Omega}_L$, $\mathbf{\Omega}_F$ are the decision spaces of the principal and the follower, respectively (see Section 2.1). We consider $PAP_{L_1F_1}$ as a causal system with endogenous observable random variables $\mathbf{V} = \{\Gamma, \Omega_L, \Omega_F, J_F, J_L\}$, where Γ is the admissible set of incentive functions with range $\mathcal{R}_\Gamma = \Gamma$. The variables Ω_L, Ω_F represent the decision variables of the principal and the follower, respectively, with ranges $\mathcal{R}_{\Omega_L} = \mathbf{\Omega}_L$ and $\mathcal{R}_{\Omega_F} = \mathbf{\Omega}_F$. The variables J_F, J_L represent the values of the utility functions of the follower and the principal, respectively.

Definition 1 *The CIDM $\mathcal{M}_{L_1F_1}$ for $PAP_{L_1F_1}$ is the CGM with endogenous variables set $\mathbf{V}$, exogenous variables set $\boldsymbol{U}$, and structural equations set $\boldsymbol{F}$ given as follows: $\mathbf{V} = \{\Gamma, \Omega_L, \Omega_F, J_F, J_L\}$; $\boldsymbol{U} = \{U_{\Omega_L}, U_{\Omega_F}, U_{J_F}, U_{J_L}\}$; and $\boldsymbol{F} = \{f_\Gamma, f_{\Omega_L}, f_{\Omega_F}, f_{J_F}, f_{J_L}\}$, with*

$$\Gamma \coloneqq \gamma; \tag{4}$$
$$\Omega_F \coloneqq f_{\Omega_F}(\gamma, \epsilon_{\Omega_F}) = BR(\gamma) + \epsilon_{\Omega_F}; \tag{5}$$
$$\Omega_L \coloneqq f_{\Omega_L}(\gamma, \Omega_F, \epsilon_{\Omega_L}) = \gamma(\Omega_F) + \epsilon_{\Omega_L}; \tag{6}$$
$$J_F \coloneqq f_{J_F}(\Omega_L, \Omega_F, \epsilon_{J_F}) = J_F(\Omega_L, \Omega_F) + \epsilon_{J_F}; \tag{7}$$
$$J_L \coloneqq f_{J_L}(\Omega_L, \Omega_F, \epsilon_{J_L}) = J_L(\Omega_L, \Omega_F) + \epsilon_{J_L}; \tag{8}$$

where $\epsilon_{\Omega_F} \in \mathcal{R}_{U_{\Omega_F}}$, $\epsilon_{\Omega_L} \in \mathcal{R}_{U_{\Omega_L}}$, $\epsilon_{\Omega_{J_F}} \in \mathcal{R}_{U_{J_F}}$, $\epsilon_{\Omega_{J_L}} \in \mathcal{R}_{U_{J_L}}$, *with* $p(\mathbf{U})$ *jointly independent. The function* $\gamma : \mathcal{R}_{\Omega_F} \to \mathcal{R}_{\Omega_L} \in \Gamma$ *is an incentive function, decided by the principal from the function space* Γ. *The functional* $BR(\gamma)$ *is the best response functional for the follower agent given* γ. *The functions* $J_F(\omega_L, \omega_F)$, $J_L(\omega_L, \omega_F)$ *are the utility functions of the follower and the leader, respectively, in the* $PAP_{L_1F_1}$.

The causal structure $\mathcal{G}_{L_1F_1}$ of $\mathcal{M}_{L_1F_1}$, has directed edges from Γ to $\{\Omega_L, \Omega_F\}$, from Ω_L to $\{J_F, J_L\}$, and from Ω_F to $\{\Omega_L, J_F, J_L\}$. The variable Γ does not depend on an exogenous random variable or any other variable in the causal system; it is a deterministic root variable whose value is decided and fixed by the principal choice of $\gamma \in \Gamma$. Thus, γ acts as a fixed input to the causal system, selected by the principal agent. The best response functional $BR(\gamma)$ in the structural Equation (5), is the mapping from the principal's chosen incentive function γ to the action ω_F^* that maximizes the follower's utility. That is,

$$BR(\gamma) = \omega_F^* \in \arg\max_{\omega_F \in \Omega_F} J_F\big(\gamma(\omega_F), \omega_F\big). \tag{9}$$

This formulation reveals how the follower's decision responds to the incentive function chosen by the principal. The structural Equation (6), $\Omega_L \coloneqq \gamma(\Omega_F) + \epsilon_{\Omega_L}$, is central to represent the inverse Stackelberg game semantics. It expresses that after the principal proposes the incentive function γ, the principal's action Ω_L is given as a function $\gamma(\omega_F)$ of the follower's action ω_F. This composition in the directed acyclic triangle $\Gamma \to \Omega_F$, $\Omega_F \to \Omega_L$, $\Gamma \to \Omega_L$, with the structural Equations (5) and (6), highlights how the principal's concrete action Ω_L responds to the follower's action, influenced by the incentive function γ selected by the principal. The utility variables J_L and J_F are defined as functions of (Ω_L, Ω_F), consistent with conventional formulations of principal–agent problems (see Section 2.1).

Building upon the CIDM $\mathcal{M}_{L_1F_1}$, Bejos et al. (2025) introduce the Causal Incentive Design (CID) Framework for causal reasoning about SS-PAPs using a historical data set $\mathcal{D} = \{\gamma^i, \omega_F^i, \omega_L^i, j_L^i\}_{i=1}^N$. Within the CID framework, the generative causal model $\mathcal{M}_{L_1F_1}$ is analyzed from the principal's perspective, where the follower's utility J_F is treated as an unobservable or hidden variable. Accordingly, the set of hidden variables for the principal is $\mathbf{H}_L = \{J_F\}$, and the set of observable variables is $\mathbf{O}_L = \{\Gamma, \Omega_L, \Omega_F, J_L\}$, with $\mathbf{H}_L \cup \mathbf{O}_L = \mathbf{V}$. Under the bilateral private information model, the principal regards the follower mechanism $f_{\Omega_F}(\gamma, U_{\Omega_F})$ in the structural equation (5) as a black box. Specifically, the principal does not know the function $BR(\gamma)$ that the follower uses to determine ω_F, but only observes the decision ω_F. This principal's perspective corresponds to a marginalization of $\mathcal{M}_{L_1F_1}$, while maintaining consistency with the bilateral private information model. `Figure 3` presents a general schematization of the CID Framework.

While Bejos et al. (2025) establish the theoretical foundations for identification and semi-parametric estimation of the causal inference target $V(\gamma) = \mathbb{E}\big[J_L \mid do(\Gamma = \gamma)\big]$, and provide regret bounds for a functional Bayesian optimization algorithm that selects a single incentive policy $\gamma^* \in \arg\max_{\gamma\in\Gamma} \hat{\mathbb{E}}\big[J_L \mid do\big(\Gamma = \gamma\big)\big]$ using a functional Gaussian process surrogate defined on a Reproducing Kernel Hilbert Space (RKHS) domain (see Hofmann et al. (2008)), the present work addresses the fundamental problem of densely sampling a CIDM $\mathcal{M}_{L_1F_1}$. The following section develops a quasi-Monte Carlo method for dense sampling of a CIDM $\mathcal{M}_{L_1F_1}$ when the infinite-dimensional intervention space variable Γ is defined by RKHS.

3. Densely Sampling a Causal Incentive Design Model

As established in previous sections, the Causal Incentive Design (CID) framework leverages historical observational data to reason about single-stage principal-agent problems (SS-PAPs). A central computational challenge in this framework is the dense sampling of a generative Causal Incentive Design Model (CIDM) like $\mathcal{M}_{L_1F_1}$, given that interventions occur over an infinite-dimensional space Γ. This section develops two algorithms to generate representative samples from a CIDM $\mathcal{M}_{L_1F_1}$ considering that the incentive function space Γ is given by a RKHS $\mathcal{H}_k$.

3.1. Finite-Dimensional Representation of the Functional Intervention

To make the evaluation and sampling of $\gamma \in \mathcal{H}_k$ computationally tractable, we employ Random Fourier Features (RFF) (Rahimi and Recht (2007)). The primary motivation of the use of RFF here is to construct a finite-dimensional mapping $z(\mathbf{x}) : \mathbb{R}^d \to \mathbb{R}^D$ that approximates the exact inner product of the infinite-dimensional feature map space of the RKHS $\mathcal{H}_k$, such that $k(\mathbf{x}, \mathbf{y}) = \langle\phi(\mathbf{x}), \phi(\mathbf{y})\rangle_{\mathcal{H}_k} \approx z(\mathbf{x})^T z(\mathbf{y})$. Using the RFF framework, the obtained D dimension approximation $z(\mathbf{x}) \approx \phi(\mathbf{x})$, with $D \in \mathbb{N}$, guaranties computational feasibility while maintaining strict convergence bounds towards the true geometry of RKHS $\mathcal{H}_k$ as $D \to \infty$.

Sampling a representative function from the feature map space relates an RKHS to the Cameron-Martin space of a Gaussian Process (GP) (see Bogachev (1998)). To make the feature map $\phi(x)$ finite and callable, the RFF framework relies on Bochner's theorem (see Rudin (1990)), which establishes that any continuous, shift-invariant kernel $k(\mathbf{x}, \mathbf{y}) = k(\mathbf{x}-\mathbf{y})$ is the Fourier transform of a positive finite measure, its spectral density $p(\omega)$. For a Gaussian kernel with bandwidth σ, the spectral density $p(\omega)$ is $\mathcal{N}(\mathbf{0}, \sigma^{-2}\mathbf{I}_d)$, where $\mathbf{I}_d$ is the $d \times d$ identity matrix. By drawing D frequencies $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_D \sim \mathcal{N}(\mathbf{0}, \sigma^{-2}\mathbf{I}_d)^D$ and a random phase shift vector $\mathbf{b} \in [0, 2\pi]_1 \times [0, 2\pi]_2 \times \cdots \times [0, 2\pi]_D$ by drawing $b_1, \ldots, b_D \sim \mathcal{U}(0, 2\pi)^D$, one can construct an explicit D-dimensional random vector map $z_{\boldsymbol{\omega},\boldsymbol{b}}(\mathbf{x}) = \sqrt{\frac{2}{D}}\cos(\boldsymbol{\omega}^T\mathbf{x} + \boldsymbol{b})$, such that its inner product approximates the exact kernel: $k(x, y) \approx z_{\omega,b}(x)^T z_{\omega,b}(y)$.

In Appendix A a Monte Carlo base procedure for sampling from a CIDM $\mathcal{M}_{L_1F_1}$ is detailed (see Algorithm 1). The extension to a quasi-Monte Carlo procedure for the sampling is described subsequently.

3.2. Quasi-Monte Carlo Extension for Sampling a CIDMs

Yang et al. (2016) extend the standard RFF approach, which relies on independent Monte Carlo (MC) sampling of spectral used in the first part of Algorithm 1, to sample a RKHS $\mathcal{H}_k$. They overcome the convergence rate of $\mathcal{O}(D^{-1/2})$ (from Rahimi and Recht (2007)) by using low-discrepancy sequences to sample spectral frequencies. We describe the use of an Owen-scrambled 1-dimensional Sobol sequence such that $\mathbf{U} \subset [0,1]$ for such a sampling in this section. This deterministic spectral mapping improves the convergence bound to approximately $\mathcal{O}(M^{-1}(\log M)^d)$, allowing for a significantly smaller approximation dimension M.

We define M as the desired number of Random Fourier Features, where M is an even integer defined by $M = 2D$. We set $D = M/2$ as the number of frequency components. The fundamental distinction of the QMC extension lies in the deterministic construction of the mapping to the spectral density. Instead of drawing pseudo-random frequencies, we generate a low-discrepancy sequence of uniformly distributed points to optimally cover the frequency space. For a one-dimensional action domain Ω_F, let $\mathbf{U} = \{u_\omega^{(j)}\}_{j=1}^D$ be a set of D points drawn from an Owen-scrambled 1-dimensional Sobol sequence such that $\mathbf{U} \subset [0,1]$. Crucially, to construct the frequencies $\boldsymbol{\omega}$, the uniform samples $u_\omega^{(j)}$ are mapped to the target normal spectral density $\mathcal{N}\left(0, \frac{1}{\sigma^2}\right)$ using the inverse normal cumulative distribution function, denoted as Φ^{-1}.

The frequency mapping is thus explicitly given by $\omega_j = \sigma^{-1}\Phi^{-1}\left(u_\omega^{(j)}\right)$. By deploying Φ^{-1}, QMC mathematically guarantees that the frequency components cover the target frequency bandwidth evenly without redundant clustering. Following the formulation in Section 4.3 for the explicit $2D$-dimensional random vector map, we utilize these D frequency components to construct the M-dimensional feature map $\mathbf{z}_{\boldsymbol{\omega}}(x) \in \mathbb{R}^M$. By employing both the sine and cosine evaluations, the need for a randomized phase shift $\mathbf{b}$ is eliminated. Specifically, the feature map is formed by concatenating the evaluations:

$$\mathbf{z}_{\boldsymbol{\omega}}(x) = \sqrt{\frac{2}{D}}\Big[\cos(\omega_1 x), \sin(\omega_1 x) \cdots \cos(\omega_D x), \sin(\omega_D x)\Big]^T.$$

To sample a function with a bounded RKHS norm, a weight vector $\mathbf{w} \in \mathbb{R}^M$ is sampled from a standard multivariate normal distribution prior $\mathbf{w} \sim \mathcal{N}(0,1)^M$. The final callable policy $\gamma(x)$ is then formulated as the linear combination $\gamma(x) = \mathbf{w}^T\mathbf{z}_{\boldsymbol{\omega}}(x)$. Having defined $\gamma(x)$, the rest of the algorithm is exactly as Algorithm 1.

4. Simulated Experiments

To empirically validate the theoretical properties of the proposed sampling methodology within the CID framework, this section details a comprehensive suite of computational experiments. These simulations assess the infinite-dimensional approximation error of the RKHS feature map (Experiment 1) and its subsequent downstream impact on the identifiability and estimation convergence of the causal inference target $\hat{V}(\gamma)$ (Experiment 2). Crucially, to address the scalability of the proposed method and move beyond simplified scalar analyses, the experiments evaluate five distinct incentive-design environments, spanning canonical non-linear geometries, classical inverse Stackelberg benchmarks, and multi-

dimensional coordination environments. The corresponding results of this experiments are summarized in Figure 1.

4.1. Canonical and Multidimensional Evaluation Instantiations

Five distinct instantiations of the single-stage model $\mathcal{M}_{L_1F_d}$ are examined.

1. Labor Economics: In this setting, the principal compensates the follower according to the level of exerted effort. Domain: $\Omega_F \in [0, 5]$. Utilities: $J_F = \Omega_L - \frac{1}{2}\Omega_F^2$ and $J_L = 3\Omega_F - \Omega_L$ (see Figure 1(b)).

2. Regulatory Economics: In this scenario, a principal regulator imposes a financial tax on a firm based on its pollution levels, thereby testing topological stability over a substantially wider domain. Domain: $\Omega_F \in [0, 10]$. Utilities: $J_F = 5\Omega_F - \Omega_L$ and $J_L = \Omega_L - \Omega_F^2$ (see Figure 1(c)).

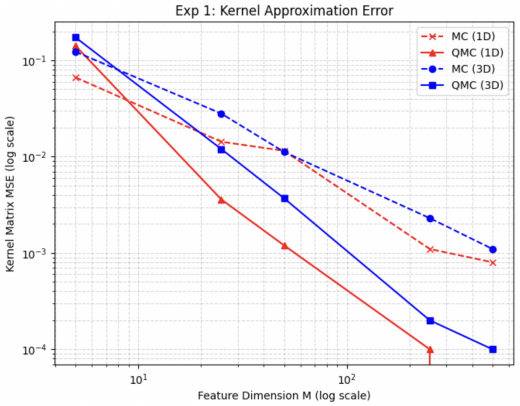
3. Exponential Platform Revenue: This scenario features a non-polynomial structure in which the follower's engagement costs increase exponentially rather than quadratically, thereby amplifying structural variance. Domain: $\Omega_F \in [0, 2]$. Utilities: $J_F = \Omega_L - e^{\Omega_F}$ and $J_L = 5\Omega_F - \Omega_L$ (see Figure 1(d)).

4. The Groot-Olsder Benchmark: This benchmark is sourced from Example 1 in Groot et al. (2014) and Example 1.1 Olsder (2009). The leader aims to achieve the global optimum at $(0, 5)$. The objective is to minimize a generic cost function for both players. Domain constraints: $\Omega_F \in [0, 7]$ and strict limits $\Omega_L \in [-1, 4]$. Costs to minimize: $J_F = \Omega_L^2 + \Omega_F^2 - \Omega_L\Omega_F$ and $J_L = (\Omega_F - 5)^2 + \Omega_L^2$ (see Figure 1(e)).

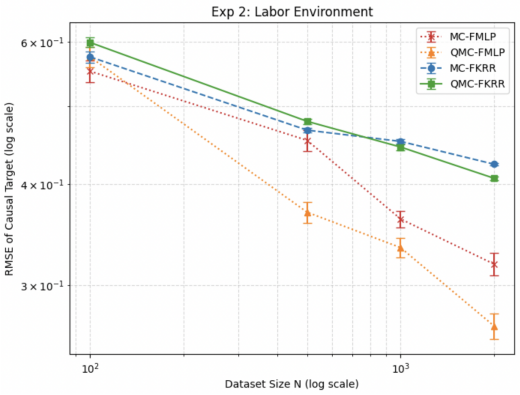
5. Multi-Task Allocation: In this setting, continuous effort is distributed across three distinct operational channels, $\mathbf{\Omega}_F \in \mathbb{R}^3$. The follower incurs independent quadratic costs for effort, while the principal obtains asymmetric linear returns from each channel. Domain: $\mathbf{\Omega}_F \in [0, 3]^3$. Utilities: $J_F = \Omega_L - \frac{1}{2}\|\mathbf{\Omega}_F\|_2^2$ and $J_L = \boldsymbol{\beta}^T\mathbf{\Omega}_F - \Omega_L$, where $\boldsymbol{\beta} = [2.0, 1.5, 3.0]^T$ (see Figure 1(f)).

4.2. Experimental Setup and Reproducibility

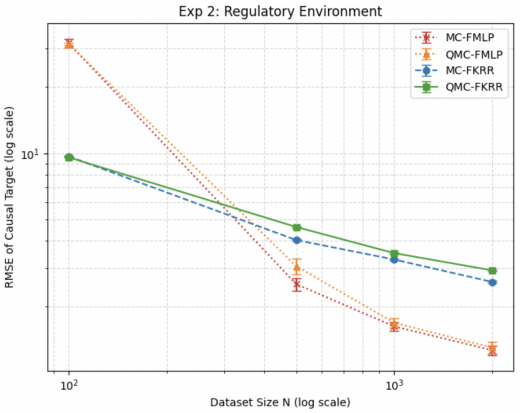
Experiment 1 (Infinite-Dimensional Approximation Error): This experiment quantifies that error by comparing the performance of standard Monte Carlo Random Fourier Features (MC-RFF) against Quasi-Monte Carlo Random Fourier Features (QMC-RFF). The primary metric for this evaluation is the Mean Squared Error (MSE) between the exact analytical kernel matrix and the approximated kernel matrix over a continuous domain. The experiment is conducted over the continuous operational action domain of the follower, such as $\Omega_F \in [0, 5]$ for the Labor Economics instantiation and $\mathbf{\Omega}_F \in [0, 3]^3$ for the Multi-Task Allocation instantiation. The exact geometry of the RKHS is defined by the underlying Gaussian Radial Basis Function (RBF) kernel with a fixed bandwidth σ. The experiment evaluates the approximation quality across a set of feature map dimensions: $M \in \{5, 25, 50, 250, 500\}$ (see Figure 1(a)).



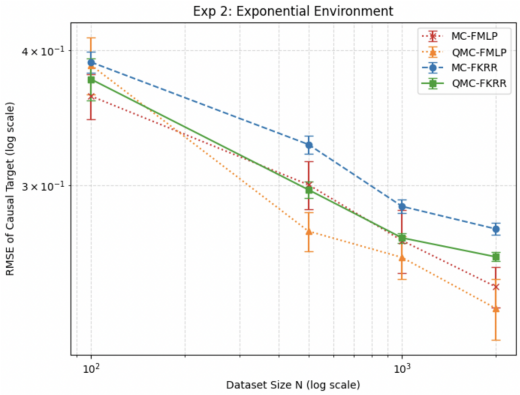
(*a*) Kernel-approximation simulations.



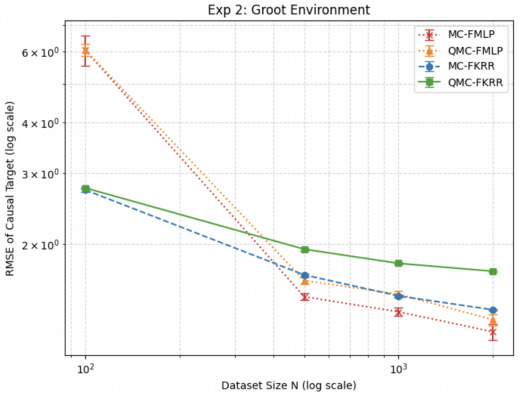
(*b*) Labor economics simulations.



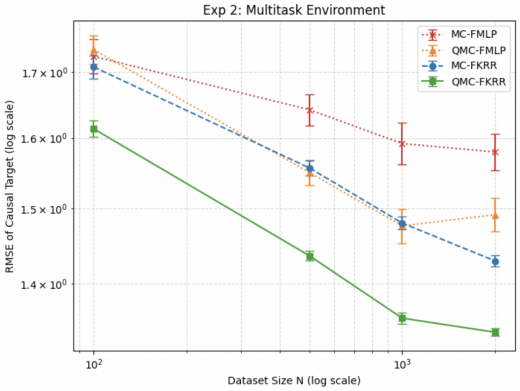
(*c*) Regulatory economics simulations.



(*d*) Exponential platform simulations.



(*e*) Groot-Olsder benchmark simulations.



(*f*) Multidimensional (3D) Multi-task allocation simulations.

Figure 1: The summary of results from the conducted experiments.

Experiment 2 (Estimation Convergence of $V(\gamma)$): This experiment is designed to evaluate how the sampling density and uniformity of the dataset—dictated by the choice of MC-RFF versus QMC-RFF algorithms—impact the downstream semi-parametric estimation of the causal inference target $\hat{V}(\gamma) = \hat{\mathbb{E}}[J_L|do(\Gamma = \gamma)]$ (Following Bejos et al. (2025)). Because the principal is operating under a bilateral private information model, they do not possess the structural equations of the follower's best-response mechanism. Consequently, they must rely on a finite observational dataset $\mathcal{D} = \{\gamma^i, \omega_F^i, \omega_L^i, j_L^i\}_{i=1}^N$ to train a surrogate model that estimates the expected utility of any proposed incentive policy.

To estimate the causal target, we apply Functional Kernel Ridge Regression (FKRR) and a Functional Multi-Layer Perceptron (FMLP) to datasets of increasing size $N \in \{100, 500, 1000, 2000\}$ generated by both sampling algorithms. Estimation accuracy is evaluated on a fixed out-of-sample hold-out test set of $M_{\text{test}} = 200$ previously unseen QMC-sampled policies. FKRR employs an RBF functional kernel mapping with $\eta = 0.01$ and Tikhonov regularization parameter $\alpha = 0.1$. The alternative FMLP architecture consists of hidden layers with 128, 64, and 32 units. To isolate the performance contributions of the proposed methodology, a complete 2×2 ablation study is conducted, crossing data generation strategies (MC versus QMC) with surrogate architectures (FMLP versus FKRR). MC-FMLP: The Functional Multi-Layer Perceptron (FMLP) neural network is trained using standard Monte Carlo (MC) sampling. This FMLP consists of three hidden layers with 128, 64, and 32 units, employs ReLU activation functions, and is optimized using the Adam algorithm. QMC-FMLP: The same FMLP architecture is applied to datasets generated via Quasi-Monte Carlo (QMC) sampling to assess whether the neural network can effectively leverage the geometrically smoothed data. MC-FKRR: The Functional Kernel Ridge Regression (FKRR) architecture is trained using standard Monte Carlo (MC) sampling. QMC-FKRR: The proposed methodology combines Quasi-Random Fourier Features with the topological RKHS structure of the FKRR surrogate.

A rigorous evaluation of estimation convergence requires establishing the exact, noise-free analytical baseline for the expected principal utility or cost under the functional interventional regime, $V(\gamma) = \mathbb{E}[J_L|do(\Gamma = \gamma)]$. Executed actions incorporate independent zero-mean Gaussian noise, represented as $\boldsymbol{\Omega}_F = \boldsymbol{\omega}_F^* + \boldsymbol{\epsilon}_{\Omega_F}$ and $\Omega_L = \gamma(\boldsymbol{\Omega}_F) + \epsilon_{\Omega_L}$, where $\epsilon \sim \mathcal{N}(0, \sigma^2)$. Because the exogenous observation and execution noises $(\epsilon_{\Omega_L}, \epsilon_{J_L})$ are independent and linearly additive in most scenarios, they vanish under expectation (see Appendix C.1). This baseline therefore isolates the structural noise of the follower's execution $(\boldsymbol{\epsilon}_{\Omega_F})$ as it propagates through the non-linear mapping.

All experiments are conducted using the following fixed hyperparameters to ensure rigorous reproducibility. Empirical metrics (MSE and RMSE) are reported as the arithmetic mean across $S = 20$ independent randomized simulation trials, each utilizing a unique random seed. Uncertainty is quantified as $\pm$ 95% confidence intervals (CIs), calculated using the standard error of the mean. The generative RKHS utilizes a Gaussian radial basis function (RBF) kernel with bandwidth $\sigma = 1.5$. Quasi-Monte Carlo (QMC) mapping is performed using an Owen-scrambled Sobol sequence, which is mapped via the inverse normal cumulative distribution function (CDF) Φ^{-1}. Structural execution noise is set to $\sigma_{\Omega_F} = 0.1$ (isotropic in multidimensional spaces), principal policy noise to $\sigma_{\Omega_L} = 0.2$, and observation noise to $\sigma_{J_L} = 0.05$.

5. Discussion

A limitation of the current approach is that unconstrained finite-feature policies may not meet ethical or economically relevant restrictions. Since the sampled interventions $\tilde{\gamma}(\mathbf{x}) = \mathbf{w}^T \mathbf{z}_{\boldsymbol{\omega}}(\mathbf{x})$ are constructed as finite linear combinations of periodic trigonometric bases, they naturally exhibit global oscillations. As a result, an unconstrained function sampled from this Gaussian process prior may readily violate practical constraints such as: boundedness (budget limitation) $\sup_{\mathbf{x}} \gamma(\mathbf{x}) \leq B_{\max}$, non-negativity: $\gamma(\mathbf{x}) \geq 0, \forall \mathbf{x} \in \Omega_F$, and monotonicity : $\nabla_{\mathbf{x}} \gamma(\mathbf{x}) \geq \mathbf{0},\ \forall \mathbf{x} \in \Omega_F$.

Although the unconstrained functions used in the preceding simulated experiments rigorously establish the topological identifiability and convergence of causal estimation for the CID framework, practical deployment sometimes requires projecting the sampled feature maps into an ethical and economically admissible subspace. Sampling functions from a continuous RKHS subject to strict global shape constraints—specifically boundedness, non-negativity, and monotonicity—presents a semi-infinite programming problem. To the end that finite-dimensional approximations reflect practical shape-constrained contracts, the policy generation algorithms can be extended through deterministic structural transformations.

To ensure that the incentive policy remains within a strict, economically valid range $[0, B_{\max}]$, the sampled Fourier expansion can be processed through a deterministic differentiable squashing function before the follower's evaluation. Denoting $\tilde{\gamma}(\mathbf{x}) = \mathbf{w}^T \mathbf{z}_{\boldsymbol{\omega}}(\mathbf{x})$ as the raw, unconstrained sampler output. The bounded, economically realistic policy may be defined using the scaled logistic sigmoid:

$$\gamma(\mathbf{x}) = \frac{B_{\max}}{1 + \exp(-\lambda \tilde{\gamma}(\mathbf{x}))},$$

where λ serves as a temperature parameter that controls the steepness of the mapping. This is a smooth, monotonic diffeomorphism that preserves the spatial coverage properties of the underlying QMC feature map. Additionally, the continuous differentiability of this squashing function integrates seamlessly with PyTorch's autograd engine, enabling constrained solvers to compute the follower's best response without disrupting the topology.

An economic contract can be rendered monotonically non-decreasing by modeling its derivative as a strictly positive unconstrained RKHS function and defining the contract as its definite integral. For a scalar action domain $\Omega_F \subset [0, x_{\max}]$, the base policy can be squared to ensure strict positivity. The monotonic intervention is then formulated as its integral from a fixed base payout $c \geq 0$:

$$\gamma(x) = c + \int_0^x \left(\mathbf{w}^T \mathbf{z}_{\boldsymbol{\omega}}(\tau)\right)^2 d\tau,$$

and therefore, the marginal incentive is precisely the squared RFF mapping: $\frac{\partial}{\partial x}\gamma(x) = (\mathbf{w}^T \mathbf{z}_{\boldsymbol{\omega}}(x))^2 \geq 0$. This ensures that the contract payout does not decrease as the follower's effort increases. Crucially, because the feature map $\mathbf{z}_{\boldsymbol{\omega}}(\tau)$ consists exclusively of sine and cosine terms, the integral of its square admits a closed-form analytical solution. This allows the monotonic contract $\gamma(x)$ to be evaluated dynamically and differentiated explicitly, eliminating the need for computationally expensive numerical quadrature during the follower's

best-response simulation. Implementing these deterministic structural transformations confines the generative causal model $\mathcal{M}_{L_1F_d}$ to the manifold of economically realistic contracts, thereby ensuring that offline dataset generation operates safely within classical constraints of applied contract theory.

6. Conclusion

This paper presents a finite-dimensional sampling framework for Causal Incentive Design Models in which incentive policies are modeled as functional interventions over an RKHS. By using Random Fourier Features, the proposed approach transforms infinite-dimensional incentive functions into explicit callable functions that can be used inside the structural equations of a CIDM. This makes it possible to simulate principal–agent interactions where the principal selects an incentive policy, the follower responds strategically, and noisy realized actions and utilities are generated. In general, the work provides a computationally tractable foundation for dense sampling of functional incentive interventions and advances the practical use of causal inference methods in principal–agent incentive design.

6.1. Future Work

To bridge the gap between computational causal inference and structural economics, future research should explicitly investigate methods for integrating economic shape constraints into the sampling process. Potential approaches include applying constrained optimization to the RFF weights, employing shape-constrained kernel regression, or substituting reproducing kernel Hilbert space (RKHS) policies with monotonic squashing functions to ensure that incentive functions are non-negative, monotonic, and budget-limited. Building on the comparisons presented in Section 4, subsequent experiments should examine other standard finite-dimensional approximation methods, such as Splines, Nyström features, or Neural Operators, to benchmark their performance against the spectral-sampling efficiency of MC-RFF and QMC-RFF. Furthermore, in online active learning contexts where data is received sequentially, extending the framework to incorporate approximations, such as sparse Gaussian processes or Nyström methods, for adaptive surrogate estimation constitutes a significant avenue for future research.

Acknowledgments

S. Bejos is supported by Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI), Scholarship Number 1562. Instituto Nacional de Astrofísica, Óptica y Electrónica.

References

Tamer Basar and Geert Jan Olsder. *Dynamic Noncooperative Game Theory*. SIAM, 2nd edition, 2009. ISBN 9780898714296.

Sebastián Bejos, Eduardo F Morales, Luis Enrique Sucar, and Enrique Munoz de Cote. Single-stage causal incentive design via optimal interventions. *Entropy*, 28(1):4, 2025.

Vladimir I. Bogachev. Infinite dimensional gaussian distributions. In *Gaussian Measures*, volume 62 of *Mathematical Surveys and Monographs*, chapter 2, pages 39–96. American Mathematical Society, Providence, RI, 1998. ISBN 978-0-8218-1054-4. doi: 10.1090/surv/062.

Sayak Ray Chowdhury and Aditya Gopalan. Kernelized bandits without regret verification. In *International Conference on Machine Learning (ICML)*, pages 832–841. PMLR, 2017.

Stephan Dempe and Alain Zemkoho. *Bilevel Optimization*, volume 161 of *Springer Optimization and Its Applications*. Springer, 2020. ISBN 978-3-030-52119-6. doi: 10.1007/978-3-030-52119-6.

Noortje Groot, Bart De Schutter, and Hans Hellendoorn. Optimal leader functions for the reverse stackelberg game: Splines and basis functions. In *2013 European Control Conference (ECC)*, pages 696–701, 2013. doi: 10.23919/ECC.2013.6669364.

Noortje Groot, Bart De Schutter, and Hans Hellendoorn. On systematic computation of optimal nonlinear solutions for the reverse stackelberg game. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 44(10):1315–1327, 2014. doi: 10.1109/TSMC.2014.2311756.

Noortje Groot, Georges Zaccour, and Bart De Schutter. Hierarchical game theory for system-optimal control: Applications of reverse stackelberg games in regulating marketing channels and traffic routing. *IEEE Control Systems Magazine*, 37(2):129–152, 2017. doi: 10.1109/MCS.2016.2643260.

Chien-Ju Ho, Aleksandrs Slivkins, and Siddharth Suri. Adaptive contract design for crowdsourcing markets: bandit algorithms for repeated principal-agent problems. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, EC '14, pages 317–334, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 978-1-4503-2567-9. doi: 10.1145/2600057.2602880.

Thomas Hofmann, Bernhard Schölkopf, and Alexander J. Smola. A review of kernel methods in machine learning. *The Annals of Statistics*, 36(3):1171–1220, 2008. doi: 10.1214/009053607000000677.

Johannes Kirschner, Mojmir Mutny, Nicolas Huesmann, and Andreas Krause. Information-directed sampling for linear bandits and beyond. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 388–396. PMLR, 2021.

Jean-Jacques Laffont and David Martimort. *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press, 2002.

David H. Mguni, Joel Jennings, Sergio Valcarcel Macua, Enrique Sison, Sofia Ceppi, and Enrique Munoz de Cote. Coordinating the crowd: Inducing desirable equilibria in non-cooperative systems. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '19, pages 386–394. International Foundation for Autonomous Agents and Multiagent Systems, 2019.

Geert Jan Olsder. Phenomena in inverse stackelberg games, part 1: Static problems. *Journal of Optimization Theory and Applications*, 143(3):555–570, 2009.

Judea Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, USA, 2nd edition, 2009. ISBN 052189560X.

Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In J. Platt, D. Koller, Y. Singer, and S. Roweis, editors, *Advances in Neural Information Processing Systems*, volume 20. Curran Associates, Inc., 2007.

Carl Edward Rasmussen and Christopher KI Williams. *Gaussian processes for machine learning*. MIT Press, 2006.

Lillian J Ratliff, Roy Dong, Shreyas Sekar, and Tanner Fiez. A perspective on incentive design: Challenges and opportunities. *Annual Review of Control, Robotics, and Autonomous Systems*, 2(1):305–338, 2019.

Alessandro Rudi and Lorenzo Rosasco. Generalization properties of random Fourier features. *Advances in Neural Information Processing Systems (NeurIPS)*, 30:3215–3225, 2017.

Walter Rudin. The basic theorems of fourier analysis. In *Fourier Analysis on Groups*, chapter 1, pages 1–34. John Wiley & Sons, Ltd, 1990. ISBN 9781118165621. doi: https://doi.org/10.1002/9781118165621.ch1.

Niranjan Srinivas, Andreas Krause, Matthias Seeger, and Sham M Kakade. Gaussian process optimization in the bandit setting: No-regret and experimental design. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 1015–1022, 2010.

Danica J Sutherland and Jeff Schneider. On the error of random Fourier features. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 862–871, 2015.

Sho Takemori, Masahiro Sato, and Taiji Suzuki. Approximation theory based methods for RKHS bandits. In *International Conference on Machine Learning (ICML)*, pages 10122–10132. PMLR, 2021.

Jiachen Yang, Ethan Wang, Rakshit Trivedi, Tuo Zhao, and Hongyuan Zha. Adaptive incentive design with multi-agent meta-gradient reinforcement learning. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '22, pages 1436–1445. International Foundation for Autonomous Agents and Multiagent Systems, 2022.

Jiyan Yang, Vikas Sindhwani, Haim Avron, and Michael W. Mahoney. Quasi-Monte Carlo feature maps for shift-invariant kernels. *Journal of Machine Learning Research*, 17(1): 4117–4154, 2016.

Appendix A. Monte Carlo Algorithm for Sampling a CIDMs

Algorithm 1 details a Monte Carlo base procedure for sampling from a CIDM $\mathcal{M}_{L_1F_1}$. Following Bochner's Theorem, we construct the explicit, finite-dimensional feature map $z_{\boldsymbol{\omega},\boldsymbol{b}}(x)$ using standard Monte Carlo (MC) sampling. Drawing D independent samples of frequencies $\omega_1, \ldots, \omega_D \sim \mathcal{N}\left(0, \frac{1}{\sigma^2}\right)^D$ and phase shifts uniformly across a full circle period $b_1, \ldots, b_D \sim \mathcal{U}(0, 2\pi)^D$ on lines 3 and 4, respectively. The randomized D-dimensional feature mapping $\mathbf{z}_{\boldsymbol{\omega},\boldsymbol{b}}(x) : \mathbb{R} \to \mathbb{R}^D$ for any input x is defined on line 5. The scaling factor $\sqrt{\frac{2}{D}}$ ensures that the inner product approximates the kernel $\mathbf{z}_{\boldsymbol{\omega},\boldsymbol{b}}(x)^T \mathbf{z}_{\boldsymbol{\omega},\boldsymbol{b}}(y) \approx k(x, y)$. To sample a function with a bounded RKHS norm $\|f\|_{\mathcal{H}_k} \leq R$, a weight vector $\mathbf{w} \in \mathbb{R}^D$ is sampled from a standard multivariate normal distribution prior $\mathbf{w} = w_1, \ldots, w_D \sim \mathcal{N}(0, 1)^D$ on line 6. A callable function sample from the RKHS is then formulated as the linear combination $f(x) = \mathbf{w}^T \mathbf{z}_{\boldsymbol{\omega},\boldsymbol{b}}(x) = \mathbf{w}^T \left(\sqrt{\frac{2}{D}} \cos(\boldsymbol{\omega}^T x + \boldsymbol{b})\right)$ on line 7. In the context of $\mathcal{M}_{L_1F_1}$, this finite-dimensional function approximation $f(x)$ acts as the incentive policy γ.

Algorithm 1: Monte Carlo Sampling of CIDM $\mathcal{M}_{L_1F_1}$

Input: Dataset size N, Kernel bandwidth σ, Approximation dimension D, Domain Ω_F, Noise distributions $p(U)$

Output: Dataset $\mathcal{D} = \left\{\gamma^i, \omega_F^i, \omega_L^i, j_L^i\right\}_{i=1}^N$

1 $\mathcal{D} \leftarrow \emptyset$

2 **for** $i \leftarrow 1$ **to** N **do**

`/* 1. Sample Callable Policy` $\gamma \in \Gamma \subseteq \mathcal{H}_k$ `*/`

3 Draw $\boldsymbol{\omega} \sim \mathcal{N}\left(0, \frac{1}{\sigma^2}\right)^D$

4 Draw $\boldsymbol{b} \sim \mathcal{U}(0, 2\pi)^D$

5 Define $\mathbf{z}_{\boldsymbol{\omega},\boldsymbol{b}}(x) := \sqrt{\frac{2}{D}} \left[\cos(\omega_1 x + b_1) \quad \cos(\omega_2 x + b_2) \quad \cdots \quad \cos(\omega_D x + b_D)\right]^T$

6 Draw $\mathbf{w} \sim \mathcal{N}(0, I_D)$

7 Define $\gamma(x) := \mathbf{w}^T \mathbf{z}_{\boldsymbol{\omega},\boldsymbol{b}}(x) = \sqrt{\frac{2}{D}} \sum_{i=1}^{D} w_i \cos(\omega_i x + b_i)$

`/* 2. Follower Best Response */`

8 Define $\mathbb{E}[J_F] := J_F(\gamma(\omega_F), \omega_F)$

9 $\omega_F^* \leftarrow \arg\max_{\omega_F \in \Omega_F} \mathbb{E}[J_F]$

`/* 3. Structural Equation Propagation */`

10 Draw exogenous noise $(\epsilon_{\Omega_F}, \epsilon_{\Omega_L}, \epsilon_{J_L}) \sim p(U)$

11 $\omega_F^i \leftarrow \omega_F^* + \epsilon_{\Omega_F}$

12 $\omega_L^i \leftarrow \gamma(\omega_F^i) + \epsilon_{\Omega_L}$

13 $j_L^i \leftarrow J_L(\omega_L^i, \omega_F^i) + \epsilon_{J_L}$

14 $\mathcal{D} \leftarrow \mathcal{D} \cup \{(\gamma, \omega_F^i, \omega_L^i, j_L^i)\}$

15 **end**

16 **return** $\mathcal{D}$

The remainder of the algorithm models a noisy Stackelberg game interaction and updates a data repository through three primary phases. In the Follower Best Response phase,

the follower defines an expected objective function and selects the optimal strategy (ω_F^*) that maximizes the expected payoff given the leader's strategy (γ). During the Structural Equation Propagation phase, stochasticity is introduced by sampling exogenous noise variables from a joint distribution $p(U)$, which are then used to perturb the follower's strategy, the leader's strategy, and the resulting leader payoff (j_L^i). In the final phase, the algorithm aggregates these realized values and the leader's policy function into a tuple, appending it to the cumulative dataset $\mathcal{D}$.

Appendix B. Related Work

B.1. Incentive design and Contract Theory

Incentive design (ID) has long been studied in economics and control through the lens of principal–agent theory Laffont and Martimort (2002); Basar and Olsder (2009). Classical approaches model uncertainty with priors over types and often assume a known mapping from actions to outcomes, producing foundational results in adverse selection and moral hazard Laffont and Martimort (2002); Ratliff et al. (2019). Control-theoretic treatments similarly emphasize leader-follower structure, but typically rely on parametric models of dynamics and payoffs (e.g., early closed-loop Stackelberg formulations and multi-level control Basar and Olsder (2009), and modern bilevel optimization theory and hardness results Dempe and Zemkoho (2020)). More recently, the machine learning community has approached ID problems using online learning and bandit formulations, where the principal adapts incentives sequentially through limited feedback Mguni et al. (2019); Ho et al. (2014); Yang et al. (2022); Ratliff et al. (2019).

B.2. RKHS Functions and Random Fourier Features

In standard Kernelized Bandits or Gaussian Process Optimization (e.g., Kernel Upper Confidence Bound or Thompson Sampling), updating the posterior distribution or calculating the objective functions scales cubically, $\mathcal{O}(t^3)$, at iteration t due to the inversion of the kernel matrix Rasmussen and Williams (2006); Srinivas et al. (2010); Chowdhury and Gopalan (2017). This severe computational bottleneck prohibits real-time or large-scale sequential optimization, motivating the development of explicit algorithmic approximations.

To resolve this computational bottleneck, the framework of Random Fourier Features (RFF) introduced by Rahimi and Recht (2007) is widely utilized to shift operations from implicit functional representations to an explicit, randomized low-dimensional subspace. Algorithmically, this structure yields a powerful Monte Carlo procedure to construct an explicit finite dimensional feature mapping $\mathbf{z}(x)$ using D random frequencies. Any target function approximation $\hat{f}(x)$ can then be represented as a linear combination over these randomized features. Sutherland and Schneider (2015) rigorously bounded the uniform approximation error of this specific cosine-phase mapping, proving that a modest feature dimension D guarantees uniform convergence over compact sets. Rudi and Rosasco (2017) expanded this theory to show that RFF approximations maintain minimax-optimal statistical learning rates in non-parametric risk minimization.

In optimization and multi-armed bandit settings, this framework does not merely serve as an approximation for static kernel matrices; it acts as a mechanism to explicitly draw

functions or *sample paths* from the RKHS. This approach has been successfully integrated into advanced sequential decision-making paradigms. Kirschner et al. (2021) utilized this explicit RFF sampling mechanism to scale Information-Directed Sampling (IDS) to non-linear landscapes, effectively mapping continuous RKHS tracking problems into manageable linear bandit subproblems. Similarly, Takemori et al. (2021) utilized these random cosine expansions to develop approximation algorithms for adversarial and stochastic RKHS bandits, establishing that tracking proxy weights instead of full history satisfies tight theoretical regret bounds.

While the standard RFF approach relies on independent Monte Carlo (MC) sampling of spectral frequencies, it suffers from a slow convergence rate of $\mathcal{O}(M^{-1/2})$, where M is the number of features. To accelerate this convergence, Yang et al. (2016) pioneered the integration of Quasi-Monte Carlo (QMC) methods with RFF. By substituting classical MC sampling with low-discrepancy sequences—specifically Halton and Sobol sequences—and mapping them through the inverse Cumulative Distribution Function (CDF) of the target spectral density, they established that the kernel integration error can be minimized at a rate approaching $\mathcal{O}(M^{-1})$. This deterministic uniform spacing ensures that the continuous function paths sampled from the underlying RKHS prior cover the spectral domain without the clusters or structural gaps inherent to random sampling.

Appendix C. Supplement to the experimental evaluation

C.1. Expected Principal Utility Ground-truth Baseline Calculation

To illustrate this analytical baseline computation of $V(\gamma)$, consider the Labor Economics scenario, where $\omega_F^* = \arg\max_{\omega_F}(\gamma(\omega_F) - 0.5\omega_F^2)$ and the principal's utility is a linear combination of the actions:

$$J_L(\Omega_L, \Omega_F) = 3\Omega_F - \Omega_L. \tag{10}$$

According to the structural equations of the CIDM, the executed actions that contain additive noise are: follower's action: $\Omega_F = \omega_F^* + \epsilon_{\Omega_F}$, principal's action: $\Omega_L = \gamma(\Omega_F) + \epsilon_{\Omega_L}$, and principal's utility: $J_L = J_L(\Omega_L, \Omega_F) + \epsilon_{J_L}$. If we substitute the actions into the utility equation, we get the fully expanded state of the world under the intervention $do(\Gamma = \gamma)$:

$$J_L = 3(\omega_F^* + \epsilon_{\Omega_F}) - \big(\gamma(\omega_F^* + \epsilon_{\Omega_F}) + \epsilon_{\Omega_L}\big) + \epsilon_{J_L}.$$

Now, we apply the expected value operator $\mathbb{E}[\cdot]$ to find the causal target $V(\gamma) = \mathbb{E}[J_L|do(\Gamma = \gamma)]$. Being a linear operator, the expectation, distributing the expectation gives:

$$V(\gamma) = 3\omega_F^* + 3\mathbb{E}[\epsilon_{\Omega_F}] - \mathbb{E}[\gamma(\omega_F^* + \epsilon_{\Omega_F})] - \mathbb{E}[\epsilon_{\Omega_L}] + \mathbb{E}[\epsilon_{J_L}].$$

Because exogenous variables are defined as independent Gaussian distributions with a mean of zero, the isolated terms disappear: $3\mathbb{E}[\epsilon_{\Omega_F}] = 0$, $\mathbb{E}[\epsilon_{\Omega_L}] = 0$, $\mathbb{E}[\epsilon_{J_L}] = 0$. The principal's execution noise (ϵ_{Ω_L}) and the environmental observation noise (ϵ_{J_L}) have absolutely no effect on the expected utility.

Notice that the term $\mathbb{E}[\gamma(\omega_F^* + \epsilon_{\Omega_F})]$ does not vanish because γ is a non-linear function (in our case, a sum of cosines from the RFF approximation) and for any non-linear function f, $\mathbb{E}[f(X)] \neq f(\mathbb{E}[X])$. Thus,

$$V(\gamma) = \mathbb{E}[J_L | do(\Gamma = \gamma)] = 3\omega_F^* - \mathbb{E}_{\epsilon_{\Omega_F} \sim \mathcal{N}(0, \sigma_{\Omega_F}^2)} [\gamma(\omega_F^* + \epsilon_{\Omega_F})] \tag{11}$$

$$= 3\omega_F^* - \int_{-\infty}^{\infty} \gamma(\omega_F^* + x) \cdot \mathcal{N}(x; 0, \sigma_{\Omega_F}^2) dx. \tag{12}$$

This expectation captures the Gaussian smoothing of the non-linear incentive function γ resulting from the follower's execution noise ϵ_{Ω_F}, providing a strict, mathematically exact baseline for verifying empirical convergence.

1. Labor Economics: The optimal follower action is $\omega_F^* = \arg\max(\gamma(\omega_F) - 0.5\omega_F^2)$. The corresponding baseline expectation is $V(\gamma) = 3\omega_F^* - \mathbb{E}_{\epsilon_{\Omega_F}}[\gamma(\omega_F^* + \epsilon_{\Omega_F})]$.

2. Regulatory Economics: The optimal follower action is given by $\omega_F^* = \arg\max(5\omega_F - \gamma(\omega_F))$. Expanding the quadratic error results in $V(\gamma) = \mathbb{E}_{\epsilon_{\Omega_F}}[\gamma(\omega_F^* + \epsilon_{\Omega_F})] - (\omega_F^*)^2 - \sigma_{\Omega_F}^2$.

3. Exponential Platform: The optimal follower action is determined by $\omega_F^* = \arg\max(\gamma(\omega_F) - e^{\omega_F})$. The corresponding baseline is $V(\gamma) = 5\omega_F^* - \mathbb{E}_{\epsilon_{\Omega_F}}[\gamma(\omega_F^* + \epsilon_{\Omega_F})]$.

4. Groot-Olsder Benchmark: The follower minimizes J_F, resulting in $\omega_F^* = \arg\min(\gamma(\omega_F)^2 + \omega_F^2 - \gamma(\omega_F)\omega_F)$. Expanding the leader's quadratic expected cost yields $V(\gamma) = \mathbb{E}_{\epsilon_{\Omega_F}}[(\omega_F^* + \epsilon_{\Omega_F} - 5)^2 + (\gamma(\omega_F^* + \epsilon_{\Omega_F}) + \epsilon_{\Omega_L})^2]$. Evaluating the expectation over independent noise terms results in $V(\gamma) = (\omega_F^* - 5)^2 + \sigma_{\Omega_F}^2 + \mathbb{E}_{\epsilon_{\Omega_F}}[\gamma(\omega_F^* + \epsilon_{\Omega_F})^2] + \sigma_{\Omega_L}^2$.

5. Multi-Task Allocation (3D): The optimal follower action is $\boldsymbol{\omega}_F^* = \arg\max(\gamma(\boldsymbol{\omega}_F) - 0.5\|\boldsymbol{\omega}_F\|_2^2)$. The baseline is $V(\gamma) = \boldsymbol{\beta}^T \boldsymbol{\omega}_F^* - \mathbb{E}_{\boldsymbol{\epsilon}_{\Omega_F}}[\gamma(\boldsymbol{\omega}_F^* + \boldsymbol{\epsilon}_{\Omega_F})]$.

C.2. Reproducible PyTorch Implementation

The experiments are implemented natively using the PyTorch ecosystem to leverage automatic differentiation and GPU acceleration. To ensure exact reproducibility and optimal QMC properties, the implementation utilizes torch.quasirandom.SobolEngine. A critical feature of this implementation is the application of Owen-type scrambling to the Sobol sequence. This preserves the strict low-discrepancy bounds crucial for eliminating the spectral clumping of Random Fourier Features, while introducing sufficient randomization to prevent linear dependencies across repeated simulation loops, ensuring statistically unbiased Monte Carlo estimates.

The exact baselines are mathematically resolved in Python via a strictly controlled two-step algorithmic pipeline.Optimization: The non-convex optimal action $\boldsymbol{\omega}_F^*$ is identified using a dense, scrambling-enabled Halton quasi-random grid search to locate the global basin, coupled natively with the SciPy L-BFGS-B constrained optimizer. We leverage PyTorch's autograd engine to calculate the exact analytical gradients dynamically fed to SciPy. The functional convolution terms $\mathbb{E}[\gamma(\dots)]$ and $\mathbb{E}[\gamma(\dots)^2]$ are strictly evaluated via high-precision numerical Monte Carlo integration. We draw $100,000$ noise samples using PyTorch tensor operations, enforcing hard domain constraints, to compute the definitive ground truth. The complete code required to execute these experiments and reproduce the results is provided in https://github.com/sebastruan/sampling-causal-incetive-design-models.

Appendix D. Supporting Figures

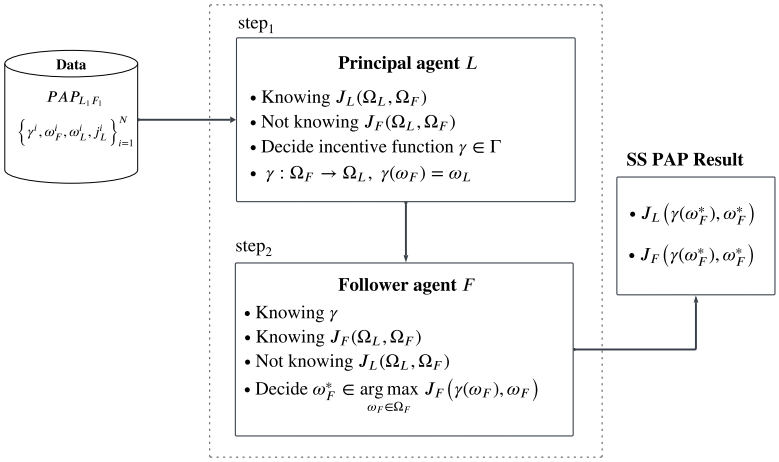


Figure 2: An schematic abstraction of a SS-PAP with historical data input.

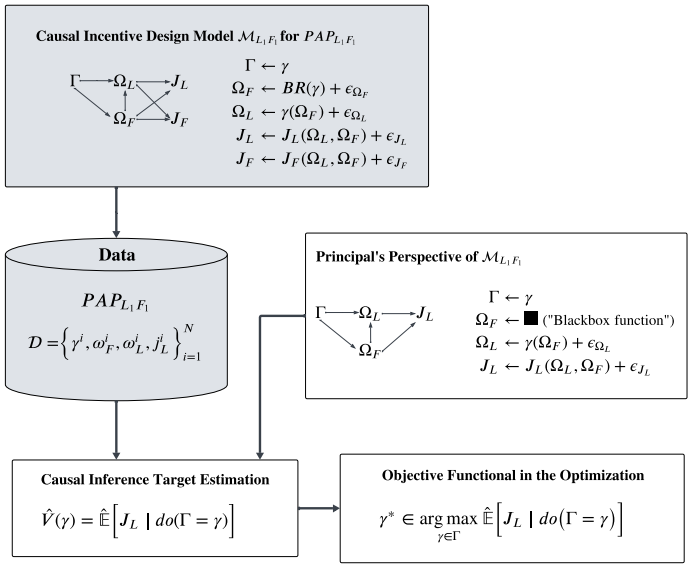


Figure 3: A general schematization of the Causal Incentive Design Framework.