观测信息引导的自适应因果实验设计

Erdun Gao[1], Liang Zhang[2,†], Jake Fawkes[3],Aoqi Zuo[4],Wenqin Liu[5], Haoxuan Li[6],Mingming Gong[5] & Dino Sejdinovic[1]
[1]澳大利亚机器学习研究所，阿德莱德大学[2]香港科技大学（广州）[3]伦敦大学学院 [4]悉尼大学[5]墨尔本大学 [6]北京大学
{erdun.gao,dino.sejdinovic}@adelaide.edu.au
liangzhang@hkust-gz.edu.cn, j.fawkes@ucl.ac.uk
aoqi.zuo@sydney.edu.au, wenqinl@student.unimelb.edu.au
hxli@stu.pku.edu.cn, mingming.gong@unimelb.edu.au

arXiv:2603.03785v1 [stat.ML] 4 Mar 2026

摘要

随机对照试验（RCTs）是因果推断的金标准，但仍是稀缺资源。尽管大规模观测数据往往可得，但它仅被用于回顾性融合，并因对偏差的顾虑而在前瞻性试验设计中被弃用。我们认为，这种“白板”数据采集策略从根本上是低效的。在本工作中，我们提出主动残差学习，这是一种将观测模型作为基础先验加以利用的新范式。该方法将实验重点从从头学习目标因果量，转向高效估计纠正观察性偏差所需的残差。为将其付诸实施，我们引入R-Design框架。理论上，我们确立了两项关键优势：（1）结构性效率差距，证明估计平滑残差对比的收敛速率严格快于重构完整结果；（2）信息效率，我们量化了标准基于参数的采集（例如BALD）中的冗余，表明此类基线将预算浪费在与任务无关的干扰不确定性上。我们提出R-EPIG（残差期望预测信息增益），这是一种直接针对因果估计目标的统一准则，可在估计中最小化残差不确定性，或为策略厘清决策边界。在合成与半合成基准上的实验表明，R-Design显著优于基线，证实修复有偏差的模型远比从头学习一个模型更为高效。

1 引言

个体治疗效应的精确估计（Wager & Athey，2018年）是从经济学（Heckman，2000年）和推荐系统（Gao 等人，2024a）到医疗保健（Foster 等人，2011年）等领域个性化决策的基础。虽然平均治疗效应（平均处理效应）只提供粗略概括，但有效的策略需要条件平均处理效应这一更细的粒度，以量化干预如何随个体而变化（Künzel 等人，2019年）。然而，准确估计条件平均处理效应从根本上受到数据二分的制约。大规模观察性研究具有总体代表性（Athey & Wager，2019年），但受隐藏混杂影响（Pearl，2009年）；而随机对照试验虽能保证无偏性，却因成本高昂和样本量有限而受到严重限制（Frieden，2017；Degtiar & Rose，2023年；Feuerriegel 等人，2024年）。这一权衡推动了关于事后数据融合的文献不断增长（Kallus 等人，2018年；Colnet 等人，2024年；Zhou 等人，2025年；Fawkes 等人，2025a）。尽管这些方法成功证明了整合两种模态的益处，但它们本质上仍是回顾性的（Song 等，2024；Cha & Lee，2025年；Zhang 等人，2025b）。它们将

†通讯作者。

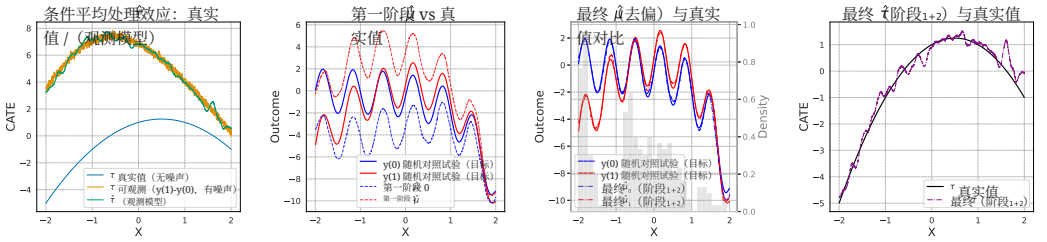


图1：R-Design在一维合成数据上的直觉。观测先验存在偏差（左），但捕捉了高频结构（中左）。将其视为固定偏移后，R-Design学习一个更简单的残差函数（中右），并利用复杂度差距以少量样本恢复真实的条件平均处理效应（右）；而朴素的观测模型则因偏差未被校正而失败。

实验数据收集视为一个固定步骤，试图在事后弥补统计功效。因此，这种被动范式未能优化数据获取过程本身，使观测先验与实验设计之间的潜在协同在很大程度上仍未被探索。

研究动机。 在许多现实应用领域中，实验之前通常已有大量观测数据（Rosenman & Owen，2021年；Epanomeritakis & Viviano，2025年）。然而，面向条件平均处理效应估计的标准因果实验设计通常采用从零开始的方法，忽略这些先验信息，从头学习因果机制（Addanki 等人，2022年；Ghadiri 等人，2023年）。我们认为，这在统计上是低效的，原因有二。首先，它忽略了由大规模观测数据所刻画的目标总体分布，而该分布刻画了因果估计目标最终被评估时所处的现实世界协变量支持集。其次，它舍弃了观测模型本身；尽管存在隐藏混杂，该模型往往仍能成功捕捉结果的全局基线结构（Kallus 等人，2018年；Hatt 等人，2022年）。这表明，改造实验设计存在一个关键机会。与其将存在混杂的观测模型视为应被丢弃的干扰项，不如将其作为基础性的信息性先验（如图1）。这一视角引出了核心研究问题：

> n **关键问题：** 我们如何利用存在混杂的观测先验来指导因果实验设计，将目标从从头探索结果转变为自适应地学习校正观察性偏差所需的内容?

挑战。 然而，将这种自适应校正策略付诸实践并非易事，面临三项根本性挑战。**1** 信号与偏差解耦。 观测先验本质上将丰富的结构信息（信号）与隐藏混杂（偏差）混杂在一起。难点在于设计一种机制，有选择地迁移全局结构知识，同时避免继承局部混杂误差，从而有效地将有用的先验形状与我们试图校正的偏差分离开来。**2** 目标错位。 识别全局观察性偏差与最大化下游性能之间存在固有差异。观测模型偏差最大的区域（误差幅度较高）可能与目标总体或决策边界并不重叠。朴素的设计可能会浪费预算，去校正无关区域中的偏差。**3**缺乏残差感知采集。 标准因果实验设计评估指标并不适用于此情形。尽管贝叶斯策略可以纳入先验，但它们通常针对完整处理效应的不确定性。因此，它们无法将预算严格集中于学习偏差校正所需的残差对比。

贡献。 为弥补这些不足，本文作出以下贡献。 **1** 一种新范式。 我们正式定义了观测信息引导的自适应因果实验设计问题。我们指出，当存在信息丰富（尽管有偏）的观测数据时，设计目标应从头重建结果曲面，根本性地转向严格识别并校正观测偏差。 **2** *R-Design*框架。我们提出一种统一的方法，包含两个核心组件：R-EPIG （第 4.1 节），一种任务感知的信息论准则，可适应估计或策略目标；以及两阶段残差（TSR）策略（第 4.2节），它在结构上将基础效应估计与不确定性量化解耦，以确保可扩展性。 **3**理论基础。 我们提供了严谨的理论分析（第5节），围绕四项关键



(a) 观测（$s = o$）　(b) 实验（$s = e$）

图 2：数据生成过程的因果图。

结果。我们正式确立了结构效率差距（引理1），证明学习残差函数在结构上比从头学习具有更高的样本效率。我们展示了目标对齐（命题 1），验证了我们的采集目标与贝叶斯条件平均处理效应风险最小化之间的等价性。尤为重要的是，我们量化了标准基线方法的信息冗余（命题 2），证明基于参数的方法会将预算浪费在干扰不确定性上。最后，我们推导出贪心策略的一致收敛速率（定理 2）。**4** 广泛的实证验证。 我们在合成与半合成基准上进行了全面实验（第 6节）。结果表明，R-Design始终优于最先进的基线方法，有力地证明了修复有偏模型比从零开始的采集效率高得多。

2 预备知识与问题设定

在本节中，我们建立本框架的数学基础。我们首先在潜在结果框架内形式化因果推断问题与 CATE 估计目标。随后，我们回顾异质数据整合所需的背景知识，以及自适应因果实验设计中所采用的特定主动学习设定。

2.1 因果估计目标与目标

潜在结果框架。我们的分析框架基于 Neyman-Rubin 潜在结果模型（Rubin，2005年；Ding，2024年）。令 $\mathrm{x} \in \mathrm{X}$ 表示协变量，$\mathrm{t} \in \{0, 1\}$ 表示二元处理分配，$\mathrm{y} \in \mathrm{Y}$ 表示观测结果。我们用 $\boldsymbol{x}$、t 和 y 表示这些变量的实现值。两个潜在结果为 y(0) 和 y(1)，分别对应于对照组和处理下的潜在结果。为适应多样的下游目标，我们定义一个一般的感兴趣目标量，记为 $\Phi(\boldsymbol{x})$，并在目标总体分布 $p_{\text{tar}}(\boldsymbol{x})$ 上评估（通常为观测总体 $p_o(\boldsymbol{x})$（Kallus 等人，2018年；ˆ Colnet 等人，2024年））。我们用帽子符号表示目标量的估计量，例如 Φ。在本文中，我们重点关注两种主要设定：

条件平均处理效应估计。 当主要目标是量化总体中的处理效应异质性时，我们以条件平均处理效应（CATE）为目标。形式上，令$\mu(\boldsymbol{x}, t) \coloneqq \mathbb{E}[\mathrm{y}(t) \mid \mathbf{x} = \boldsymbol{x}]$ 表示处理组 t 的期望潜在结果曲面。CATE 定义为处理组与对照组响应曲面之间的逐点差：

$$\tau(\boldsymbol{x}) \coloneqq \mu(\boldsymbol{x}, 1) - \mu(\boldsymbol{x}, 0). \tag{1}$$

该目标估计量刻画了处理效应如何随协变量 $\boldsymbol{x}$ 的函数而变化。估计质量由异质效应估计精度（PEHE）（Hill，2011年）衡量，它量化了估计量在目标分布上的期望平方误差（或 L_2 风险）：

$$\epsilon_{\text{PEHE}}(\hat{\tau}) = \mathbb{E}_{\mathbf{x} \sim p_{\text{tar}}(\cdot)}\left[(\hat{\tau}(\mathbf{x}) - \tau(\mathbf{x}))^2\right]. \tag{2}$$

决策。 或者，目标也可以是策略优化，即确定每个个体的最优处理分配，以最大化总体福利（Fernández-Loría 和 Provost，2022年）。这等同于识别处理效应的符号，而非估计其连续数值。目标估计量是二元最优策略 $\Phi(\boldsymbol{x}) = \pi(\boldsymbol{x}) = \mathbb{I}(\tau(\boldsymbol{x}) > 0)$，它指定能够带来较高期望结果的处理组。与条件平均处理效应估计不同，这项任务意味着只有在.....附近才需要较高精度

决策边界（其中 $\tau(\boldsymbol{x}) \approx 0$）。绩效通过平均策略误差（APE）衡量（Gao 等人，2024b），其定义为在目标总体中分配到次优处理的概率（误分类率）：

$$\epsilon_{\mathrm{APE}}(\hat{\pi}) = \mathbb{E}_{\mathbf{x}\sim p_{\mathrm{tar}}(\cdot)}\left[\mathbb{I}(\hat{\pi}(\mathbf{x}) \neq \pi(\mathbf{x}))\right]. \tag{3}$$

2.2 异质数据来源

我们考虑一个涉及两个不同数据来源的场景：一个大型观察性研究（OS）和一个较小的实验数据集，如图 2 所示。我们用随机变量 $\mathrm{s} \in \{o, e\}$ 表示数据来源。对于每个数据来源 s，我们将来源特定条件均值定义为 $\mu_s(\boldsymbol{x}, t) \coloneqq \mathbb{E}[\mathrm{y} \mid \mathrm{x} = \boldsymbol{x}, \mathrm{t} = t, \mathrm{s} = s]$。来源 s 的来源特定对比为 $\tau_s(\boldsymbol{x}) \coloneqq \mu_s(\boldsymbol{x}, 1) - \mu_s(\boldsymbol{x}, 0)$。该量衡量观测到的条件均值差。对于 OS，由于存在隐藏混杂，对比 $\tau_o(\boldsymbol{x})$ 通常有偏（Fawkes 等人，2025a）。对于实验数据来源，受控分配机制确保 $\tau_e(\boldsymbol{x})$ 能够在实验支撑集 $\mathrm{X}_e \subseteq \mathrm{X}$ 内识别真实的条件平均处理效应（CATE）：

$$\tau(\boldsymbol{x}) = \tau_e(\boldsymbol{x}), \quad \text{for } \boldsymbol{x} \in \mathcal{X}_e. \tag{4}$$

为整合这些不同来源，我们通过以下分解，刻画有偏的观测对比 $\tau_o(\boldsymbol{x})$ 与真实因果效应 $\tau(\boldsymbol{x})$ 之间的结构关系：

$$\tau(\boldsymbol{x}) = \tau_o(\boldsymbol{x}) + \tau_\delta(\boldsymbol{x}), \quad \text{where } \tau_\delta(\boldsymbol{x}) \coloneqq \delta(\boldsymbol{x}, 1) - \delta(\boldsymbol{x}, 0). \tag{5}$$

这里，$\tau_\delta(\boldsymbol{x})$ 表示残差对比（去偏校正），用于捕捉观察性相关与潜在因果机制之间的差异。为确保该分解能够得到有效且可识别的因果目标，我们对数据生成机制施加以下假设。

假设 *1*。稳定单位处理值假设：$y = ty(1) + (1 - t)y(0)$，且单位之间不存在相互干扰。设计无混杂性：实验中的处理分配由设计策略控制，且仅取决于观测信息，从而确保 $t \perp \{y(0), y(1)\} \mid (x, s = e)$。设计空间正性：实验设计允许在关注区域内探索两个处理臂，$0 < p(t = 1 \mid x, s = e) < 1$。效应不变性（可迁移性）：潜在因果机制在不同来源之间保持不变，即 $\forall s \in \{o, e\}\ \mathbb{E}[y(1) - y(0) \mid x = \boldsymbol{x}, s = s] = \tau(\boldsymbol{x})$。这排除了 *CATE* 发生非自然偏移的情况，即实验方案本身而非处理干预了结果机制。支撑包含：目标总体的支撑集包含于可行实验区域内，$supp(p_{tar}) \subseteq X_e$。

假设 1 确保 $\tau_\delta(\boldsymbol{x})$ 是一个结构一致的量，表示纯粹的观察性偏差。可迁移性保证 $\tau_\delta(\boldsymbol{x})$ 捕捉的是混杂机制，而非底层条件平均处理效应中的偏移，从而使因果目标可通过残差校正加以识别。

2.3 问题设定

与数据收集固定的回顾性融合不同，我们研究有限实验预算 n_B的动态分配。我们假设可以访问一个静态的、带标签的观察性数据集$\mathrm{D}_O = \{(\boldsymbol{x}_i, t_i, y_i, s_i = o)\}_{i=1}^{n_O}$。我们的目标是针对由边际协变量分布 $p_{\mathrm{tar}}(\boldsymbol{x})$刻画的目标总体优化条件平均处理效应估计量。按照标准惯例，我们假设观测协变量能够代表该目标（即D_O 的经验分布近似于 p_{tar}）（Kallus 等人，2018年；Colnet 等人，2024年；Yang 等人，2025年）。在实验阶段，我们拥有一个由可招募单元组成的有限候选池$\mathrm{D}_P = \{\boldsymbol{x}_j\}_{j=1}^{n_P}$。实验数据集$\mathrm{D}_E$最初为空。设计过程按顺序阶段 $k = 1, \ldots\ldots, n_B$进行。在每个阶段：1 学习器联合选择一个单元 $\boldsymbol{x}_k \in \mathrm{D}_P$ 和一个处理 $t_k \in \{0, 1\}$；2 学习器观测实验结果 $\ldots$；3 数据集更新为：$\mathrm{D}_E^{(k)} \leftarrow \mathrm{D}_E^{(k-1)} \cup \{(\boldsymbol{x}_k, t_k, y_k, \mathrm{s}_k = e)\}$，并将该单元从候选池$\mathrm{D}_P$中移除。

> n **设计目标**。 这项任务的核心挑战是设计一个有原则的效用函数，$U(\cdot \mid \mathrm{D}_O, \mathrm{D}_E^{(k-1)})$，用于量化特定查询（$\boldsymbol{x}$， t）所带来的信息增益。采集

策略在阶段 k 中，于联合乘积空间上最大化该效用：

$$(\boldsymbol{x}_k^*, t_k^*) = \underset{(\boldsymbol{x},t)\in\mathcal{D}_P\times\{0,1\}}{\arg\max} \; U(\boldsymbol{x}, t \mid \mathcal{D}_O, \mathcal{D}_E^{(k-1)}). \tag{6}$$

3 观测信息引导的因果实验设计原则

从结果探索到残差校正。 标准实验设计通常采用白板 式方法，耗费预算从头学习完整的结果曲面。我们认为，这在结构上效率低下。给定分解 $\tau(\boldsymbol{x}) = \tau_o(\boldsymbol{x}) + \tau_\delta(\boldsymbol{x})$，其中观测对比 $\tau_o(\boldsymbol{x})$ 在实验开始前已估计，若将 τ_o 视为固定偏移，则总效应的条件熵等同于残差的条件熵：$\mathrm{H}(\tau(\boldsymbol{x}) \mid \mathrm{D}_E) \equiv \mathrm{H}(\tau_\delta(\boldsymbol{x}) \mid \mathrm{D}_E)$。因此，任何重新学习预估信号 $\tau_o(\boldsymbol{x})$ 的采集都不会带来边际效用。这种结构上的等价关系决定了我们的第一条原则：

原则 1（残差不确定性原理）。 实验设计应针对残差对比 τ_δ（$\boldsymbol{x}$）的认知不确定性。采集函数应将估计得到的观测对比 τ_o（$\boldsymbol{x}$）视为固定均值函数，利用预算消除有偏先验与干预真值之间的结构性差异，而非重新学习完整的结果曲面。

从全局探索到目标估计量对齐。 虽然原则 1 确定了结构性目标（残差），但并未说明预算的空间分配。经典主动学习准则通常旨在全局最小化模型方差，这隐含地假设整个领域具有均匀的效用（Houlsby 等人，2011年）。然而，因果推断本质上以目标为导向；其有效性完全取决于特定的目标估计量（例如 CATE 或策略）和目标总体p_{tar}（$\boldsymbol{x}$）（Hernán 和 Robins；Smith 等人，2023年）。因此，在与目标密度无关或对决策边界无关紧要的区域降低残差不确定性，几乎没有价值。纯粹的认知不确定性并不等同于效用。

原则 2（目标估计量对齐采集）。最优实验设计必须以风险为中心，而非以不确定性为中心。设计点（$\boldsymbol{x}, t$）的效用并非内在固有，而仅由其对目标特定风险的预期降低幅度来衡量。这意味着实验预算必须非均匀地集中于残差不确定性较高、且对指定目标估计量影响重大的区域。

4 THER-DESIGN 框架

本节介绍 R-Design，这是一种将因果实验设计转化为主动残差学习的框架。借助分解式 $\tau(\boldsymbol{x}) = \hat{\tau}_o(\boldsymbol{x}) + \tau_\delta(\boldsymbol{x})$，其中观测对比 $\hat{\tau}_o$ 充当预计算偏移量，R-Design 明确以识别残差 τ_δ 为目标。我们首先推导基于残差的采集准则 R-EPIG，然后介绍主动学习（AL）策略，以实现对该准则的可处理计算。

4.1 治疗设计准则：R-EPIG

我们的框架通过明确地最小化特定下游目标 $\Phi(\cdot)$ 的认知不确定性，将原则 2 落实为实际操作。关键在于，由于观测基线 $\hat{\mu}_o(\boldsymbol{x}, t)$ 被视为固定均值函数，实验查询 $(\boldsymbol{x}, t)$ 的信息量完全来自残差结果，记作随机变量 $\mathrm{r} := \mathrm{y} - \hat{\mu}_o(\boldsymbol{x}, t)$。我们提出残差预期预测信息增益（R-EPIG）。不同于基于原始结果的标准准则（Smith 等，2023），R-EPIG 通过观测残差 r，专门选择能够最大化关于目标 $\Phi(\cdot)$ 的期望信息增益的实验。令 H_k 表示步长 k 时的历史信息。一般的采集目标为：

$$\alpha_{\text{R-EPIG}}(\boldsymbol{x},t) \coloneqq \mathbb{E}_{\mathbf{x}^*\sim p_{\text{tar}}(\cdot)}\Big[\mathrm{I}\big(\mathrm{r};\,\Phi(\mathbf{x}^*)\mid \boldsymbol{x},t,\mathcal{H}_k\big)\Big]. \tag{7}$$

这里，目标 $\Phi(\cdot)$ 的随机性完全由残差过程的后验分布引起。外层期望则对目标总体上的这一增益取平均。

注 1. 虽然 $\mathrm{I}(\mathrm{r};\Phi) \equiv \mathrm{I}(\mathrm{y};\Phi)$，但将目标表述为关于 r 的形式，与我们的残差高斯过程相一致。这在结构上将推断与 $\mathcal{D}_O$解耦，从而确保复杂度仅随实验预算 n_B 而非观测样本规模 n_O增长。

通过指定 $\Phi(\cdot)$，该框架得到两种面向任务的策略：*R-EPIG-Est* 用于估计，*R-EPIG-Policy*用于策略。

4.1.1 面向估计的设计（R-EPIG-EST）

当目标是最小化 PEHE 时，我们针对校正项的幅度。我们提出两种设定：

联合结果残差（R-EPIG-μ）。为恢复完整的潜在结果曲面，我们将目标定义为残差向量 $\Phi(\boldsymbol{x}) = \boldsymbol{\delta}(\boldsymbol{x}) \coloneqq [\delta(\boldsymbol{x},0)\ \delta(\boldsymbol{x},1)]^\top$。由于观测基线 $\hat{}\mu_o$被视为固定偏移，降低 $\boldsymbol{\delta}(\boldsymbol{x})$ 中的不确定性会直接降低绝对潜在结果 $\mu_e(\boldsymbol{x},t)$ 中的不确定性。该准则为：

$$\alpha^{\mu}_{\text{R-EPIG}}(\boldsymbol{x},t) \coloneqq \mathbb{E}_{\mathbf{x}^*\sim p_{\text{tar}}(\cdot)}\Big[\mathrm{I}\big(\mathrm{r};\,\boldsymbol{\delta}(\mathbf{x}^*)\mid \boldsymbol{x},t,\mathcal{H}_k\big)\Big]. \tag{8}$$

该策略虽保守但稳健：它识别出观测先验对任一处理组均不可靠的区域。

直接条件平均处理效应残差（R-EPIG-τ）。为了学习条件平均处理效应，我们可以将目标严格定义为残差对比 $\Phi(\boldsymbol{x}) = \tau_\delta(\boldsymbol{x}) \coloneqq \delta(\boldsymbol{x},1) - \delta(\boldsymbol{x},0)$。该准则为：

$$\alpha^{\tau}_{\text{R-EPIG}}(\boldsymbol{x},t) \coloneqq \mathbb{E}_{\mathbf{x}^*\sim p_{\text{tar}}(\cdot)}\Big[\mathrm{I}\big(\mathrm{r};\,\tau_\delta(\mathbf{x}^*)\mid \boldsymbol{x},t,\mathcal{H}_k\big)\Big]. \tag{9}$$

该变体在理论上样本效率更高。通过以差值为目标，它会自然忽略在相减中相互抵消的共享基线误差，从而避免将预算浪费在与处理效应无关的变异上。

4.1.2 面向策略的设计（R-EPIG-POLICY）

在最小化APE时，目标估计量退化为二元决策边界。我们将目标定义为策略指示函数 $\Phi(\boldsymbol{x}) = \pi(\boldsymbol{x}) \coloneqq \mathbb{I}(\hat{}\tau_o(\boldsymbol{x}) + \tau_\delta(\boldsymbol{x}) > 0)$。该准则针对关于该符号的信息增益：

$$\alpha^{\pi}_{\text{R-EPIG}}(\boldsymbol{x},t) \coloneqq \mathbb{E}_{\mathbf{x}^*\sim p_{\text{tar}}(\cdot)}\Big[\mathrm{I}\big(\mathrm{r};\,\pi(\mathbf{x}^*)\mid \boldsymbol{x},t,\mathcal{H}_k\big)\Big]. \tag{10}$$

与方差缩减不同，R-EPIG-Policy是决策感知的：它会忽略总效应的符号已能以高置信度确定的区域，将预算严格集中于解决决策边界附近的模糊性。

注释 2（信息增益与直接误差降低）。 虽然最小化 *APE* 是最终目标，但直接误差降低面临信号稀疏的问题：对于远离当前边界的候选项，改变策略的概率微乎其微，导致效用为零的平台。相比之下，π 上的熵减是一种平滑的替代指标，能够提供有用的指引，引导学习器走向边界，即使此时发生符号翻转的可能性仍然很低。

4.2 模型架构：THETSR 策略

为在 R-EPIG 中具体化概率项，我们引入两阶段回归（TSR）策略。TSR 在结构上将容量需求与不确定性 需求分离。

阶段 1：观测基础学习器。我们首先利用丰富的观测数据 D_O，学习潜在结果的高容量基线估计。这个阶段与模型无关，因此可以

使用最先进的估计量$\mathcal{M}_O$（例如，TabPFN（Hollmann 等人，2025年）、CausalPFN（Bal- azadeh 等人，2025年））来捕捉复杂的全局结构信号。由此得到的结果曲面$\hat{}\mu_o(\boldsymbol{x}, t)$被视为预先计算的函数偏移。该基线模型充当后续残差模型的固定均值函数，使主动学习器无需重建全局结果曲面，从而能够将有限的容量用于修正局部偏差。

阶段 2：贝叶斯残差学习。序贯阶段严格地对残差向量的认知不确定性进行建模$\boldsymbol{\delta}(\boldsymbol{x})$。学习$\boldsymbol{\delta}(\boldsymbol{x})$面临根本性的反事实挑战：对于任何个体，我们只能观测到事实残差$r_k = y_k - \hat{}\mu_o(\boldsymbol{x}_k, t_k)$。为解决这一问题，我们采用多任务高斯过程（MTGP）（Bonilla 等人，2007）作为后端推理引擎。我们选择 MTGP，是因为其核结构显式编码了因果相关，使得信息能够在不同处理臂之间共享，从而推断缺失的反事实残差：

$$\boldsymbol{\delta}(\boldsymbol{x}) \sim \mathcal{GP}(\mathbf{0}, \boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x}')), \quad r \mid \boldsymbol{x}, t \sim \mathcal{N}(\delta(\boldsymbol{x}, t), \sigma^2). \tag{11}$$

通过将核$\boldsymbol{K}$分解为任务特异和输入特异分量（例如，通过 CMGP（Alaa 和 Van Der Schaar，2017）、NSGP（Alaa 和 Schaar，2018年）），模型能够同时捕捉偏差的光滑度以及潜在结果残差之间的相关性。可扩展性。标准高斯过程的计算复杂度呈立方增长，$\mathcal{O}(n^3)$。TSR 解决了这一瓶颈：高斯过程仅拟合实验集中的残差，而该实验集规模较小$\mathcal{D}_E$。复杂度随预算$n_E \ll n_O$增长，而非随庞大的观测样本规模增长，从而确保序贯采集能够实时进行。

注 3（结构效率的直觉）。*TSR* 的效率依赖于一个关键的结构性洞见：完整的结果机制（例如生物反应）通常复杂且需要大量数据，而混杂偏倚（例如临床指南）在总体中的变化通常更加平滑。*TSR* 将复杂的基线学习交由$\mathcal{D}_O$完成，仅将昂贵的实验预算用于这个更简单的残差，从而加快收敛。

4.3 预算约束采集算法

R-EPIG 效用驱动着我们在固定预算下进行条件平均处理效应估计的主动学习策略（算法 1）。该过程采用两阶段架构。首先，在$\mathcal{D}_O$上训练观测模型，并将其冻结，作为函数偏移。其次，在自适应残差学习阶段，我们迭代地选择效用高的点组成批次。与贪心方法不同，我们根据效用得分进行 softmax 加权采样，以确保采集批次具有多样性。查询实验结果后，立即通过减去冻结的函数偏移将其转换为残差，并用于更新$\hat{}$残差模型$\boldsymbol{\delta}$。这一循环持续至预算耗尽，最终得到的估计量将学习到的残差对比叠加到观测基线上。

5 理论分析

我们通过确立四大理论支柱，为所提出的 R-Design 框架提供理论依据。首先，我们指出一项根本性的 结构优势 (引理1)，证明残差对比的极小极大收敛速率严格快于完整结果曲面。其次，为充分利用这一优势，我们推导出一项 目标对齐 结果（命题 1），验证了最小化残差不确定性在数学上等价于最小化贝叶斯 PEHE 风险。第三，我们展示了该方法的信息效率 （命题 2），正式证明基于参数的基线方法（如 BALD）优化的是一个宽松的界，其中包含与任务无关的干扰信息，而 R-EPIG 则直接针对目标估计量。最后，我们给出一项 算法不确定性收敛保证（定理 2），表明我们的贪心策略能够有效实现复杂度分析所预测的加速收敛速率。

5.1 样本复杂度：结构效率差距

我们通过霍尔德空间上的极小极大收敛速率来量化 R-Design 的效率（Yang& Tokdar，2015）。采用关于 Z = X × {0,1} 的联合函数视角，我们将真实的潜在结果分解为固定的观测基线与可学习残差：$\mu_e(\boldsymbol{x}, t) = \mu_o(\boldsymbol{x}, t) + \delta(\boldsymbol{x}, t)$。其中，$\mu_o$捕捉可通过观测数据获取的相关性，而$\delta$表示恢复$\mu_e$所需的因果校正。令$\alpha_{\cdot,t}$和$d_{\cdot,t}$分别表示处理组$t$的光滑度和有效维数。

引理1。 假设观测基础学习器 $\hat{}\mu_o$ 和实验残差学习器 δ 采用的先验分别针对其目标函数的光滑度进行了最优适配。在标准假设（重叠性和实验无混杂性）下，期望 *PEHE* 风险 $\psi(\hat{}\tau)$满足以下上界：

$$\psi(\hat{\tau}) \leq C_1 \underbrace{\max_{t\in\{0,1\}} n_E^{-\frac{2\alpha_{\delta,t}}{2\alpha_{\delta,t}+d_{\delta,t}}}}_{\textit{Residual Convergence}} + C_2 \underbrace{\max_{t\in\{0,1\}} n_O^{-\frac{2\alpha_{\mu_o,t}}{2\alpha_{\mu_o,t}+d_{\mu_o,t}}}}_{\textit{Observational Baseline}} + \epsilon_{\text{approx}}, \tag{12}$$

其中， C_1、 $C_2 > 0$ 是依赖于定义域的常数，ϵ_{approx} 表示不可约近似误差。

解读： 引理1 通过残差复杂度决定风险。机制主导性：由于$n_E \ll n_O$，观测误差相当于一个静态基线，使收敛主要由第一项主导。效率差距：标准的"白板"方法受限于原始结果$^{-\frac{2\alpha_{\mu_e}}{}}$复杂度$\mathcal{O}(n_E^{2\alpha_{\mu_e}+d_{\mu_e}})$。相比之下，R-Design 利用了残差通常更平滑这一结构性质（$\alpha_\delta > \alpha_{\mu_e}$），从而实现更快的衰减率并降低样本复杂度。详细证明见附录 Alaa 和 Schaar(2018年，定理2）中的界经调整以适用于我们的残差分解后的结果。附录 D.4。

5.2 目标：残差学习的最优性

引理 1 表明，残差在理论上是更优的学习目标。现在我们将说明，明确以降低残差不确定性为目标，是最小化因果误差的有效替代指标。由于观测对比 $\hat{}\tau_o(\boldsymbol{x})$ 在给定 $\mathcal{D}_O$的条件下是确定的，因此 $\tau = \hat{}\tau_o + \tau_\delta$ 的后验不确定性与残差对比的后验不确定性相同： $\mathbb{V}[\tau \mid \mathcal{D}_E,\ \mathcal{D}_O] \equiv \mathbb{V}[\tau_\delta \mid \mathcal{D}_E]$。

命题 1 （目标对齐）。 假设条件平均处理效应估计量是后验均值 $\hat{}\tau_k(\boldsymbol{x}) = \mathbb{E}[\tau(\boldsymbol{x}) \mid \mathcal{H}_{k-1}]$。对于以总体 p_{tar} 为目标的主动条件平均处理效应估计，在 ***R-Design*** 分解下，最小化期望贝叶斯*PEHE*风险的最优选择可简化为最小化残差对比 $\tau_\delta(\boldsymbol{x})$的积分后验方差：

$$\begin{aligned}(\boldsymbol{x}_k^\dagger, t_k^\dagger) &= \arg\min_{(\boldsymbol{x},t)} \mathbb{E}_y\Big[\mathbb{E}_{\mathbf{x}^*\sim p_{\text{tar}}}\big[(\hat{\tau}_k(\mathbf{x}^*) - \tau(\mathbf{x}^*))^2 \mid \mathcal{D}_+\big]\Big] \\ &\equiv \arg\min_{(\boldsymbol{x},t)} \mathbb{E}_y\Big[\int \mathbb{V}[\tau_\delta(\boldsymbol{x}^*) \mid \mathcal{D}_+]\, dp_{\text{tar}}(\boldsymbol{x}^*)\Big],\end{aligned} \tag{13}$$

其中，$\mathcal{D}_+ = \mathcal{D}_E^{(k-1)} \cup \{(\boldsymbol{x}, t, y)\}$ 是更新后的历史数据，外层期望是关于预测后验分布 $p(y \mid \boldsymbol{x}, t, \mathcal{H}_{k-1})$ 求取的。

实用替代指标。 虽然式 13（积分方差缩减）是黄金标准，但其计算在实际中难以处理。相反，R-Design 最大化信息增益 I（r；τ_δ），将其作为计算高效的替代指标。在高斯过程中，熵减与方差最小化之间具有单调性，因此这一松弛是合理的（见附录 D.3.2）。

5.3 超越基于参数的采集

鉴于学习残差所具有的结构优势（引理 1），关于采集策略自然会产生一个问题：标准的基于参数的主动学习方法（例如，以残差模型参数 θ为目标的 BALD）能否实现最优效率？我们通过量化参数学习与估计目标学习之间的信息论差距，给出了否定的答案。

命题 2 (信息冗余)。 设 θ 为残差模型的参数， $\tau_\delta(\boldsymbol{x}^*)$为目标 $\boldsymbol{x}^*$处的残差对比。参数导向效用（*Residual-BALD*）可分解为目标效用（*R-EPIG*）与一个非负冗余项之和：

$$\underbrace{\mathrm{I}(y;\theta \mid \mathcal{D})}_{\textit{Residual-BALD}} = \underbrace{\mathrm{I}(y;\tau_\delta(\boldsymbol{x}^*) \mid \mathcal{D})}_{\textit{Target Utility (R-EPIG)}} + \underbrace{\mathrm{I}(y;\theta \mid \tau_\delta(\boldsymbol{x}^*), \mathcal{D})}_{\textit{Nuisance Redundancy}}. \tag{14}$$

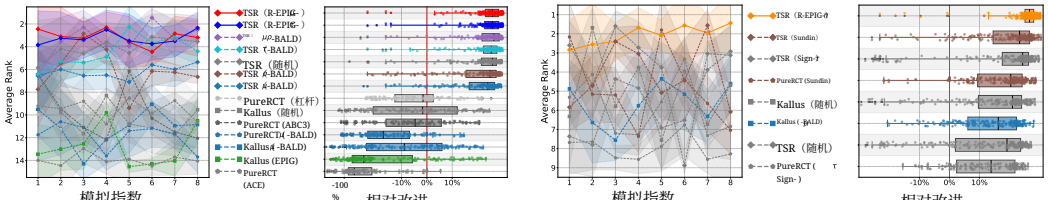


图3：使用平均排名指标和相对性能提升，对效应估计和策略学习任务中的采集函数进行比较；这些任务涵盖基础效应函数、偏差函数和CATE函数的八种组合。

过度建模的代价。 项 $\Delta = \mathrm{I}(y;\theta \mid \tau_\delta(\boldsymbol{x}^*), \mathcal{D})$ 表示 干扰信息：获取的有关模型内部结构的信息位（例如高频基函数系数或权重置换），在计算对比 τ_δ时，这些信息实际上会相互抵消。通过最大化左侧项，基于参数的方法不可避免地会将部分实验预算浪费在减少这种无关不确定性上。相比之下，R-EPIG 直接针对第一项，确保每个获取的样本都只用于改进目标总体上的因果估计目标。我们在附录 D.5中基于互信息链式法则给出了严格证明。

5.4 信息论收敛

在确立了 R-EPIG 目标函数的效率后，我们最后分析它是否实现了引理 1 预测的加速收敛。我们采用将误差降低与最大信息容量联系起来的传导式主动学习界 $\gamma_n(\cdot)$ (Hübotter 等人，2024年)。

定理2 （不确定性收敛）。 令 $\gamma_{n_B}(\tau_\delta)$ 表示残差对比函数的最大信息容量。在贪心 *R-EPIG* 策略下（假设满足 α-弱次模性），残差对比的后验方差一致收敛。具体而言，以高概率，对所有目标点 $\boldsymbol{x}^* \in \mathcal{X}$，均有：

$$\mathbb{V}[\tau_\delta(\boldsymbol{x}^*) \mid \mathcal{D}_{n_B}] \le \eta^2_{\text{feas}}(\boldsymbol{x}^*) + \tilde{\mathcal{O}}\left(\frac{\gamma_{n_B}(\tau_\delta)}{\sqrt{n_B}}\right), \tag{15}$$

其中， $\eta^2_{\text{feas}}(\cdot)$ 表示可行不可约不确定性。由于收敛速率项仅取决于 n_B和容量，因此这意味着一致收敛至不可约不确定性下限。

效率差距。 以完整曲面为目标的标准主动学习器 τ 满足由更大容量 $\gamma_{n_B}(\tau)$所决定的相同界。关键在于，由于 τ和 τ_δ具有完全相同的不可约不确定性下限（$\mathbb{V}[\tau|\mathcal{D}] \equiv \mathbb{V}[\tau_\delta|\mathcal{D}]$），性能差距完全由容量比决定。正如引理 1所示，由于 $\alpha_\delta > \alpha_{\mu_e}$，残余容量 $\gamma_{n_B}(\tau_\delta)$ 的增长速度显著低于完整结果容量 $\gamma_{n_B}(\tau)$（例如，对数级与多项式级）。从结构上保证了 R-Design 能实现更优的收敛速率。详细证明见附录 D.3，其中通过谱衰减证明一致收敛，并处理近似次模性。

6 实验结果

在本节中，我们报告实证结果，以验证我们的框架在获取数据点方面的有效性，涵盖多个合成数据集以及两个半合成数据集，具体为婴儿健康与发展计划（IHDP）（Hill，2011年）和艾滋病临床试验组第 175 项研究（ACTG-175）（Hammer 等人，1996年）。所有数据集的详细信息见附录 B.3。

基线采集函数。 我们将本方法与一组多样化的采集策略进行基准比较，涵盖从随机采样到先进基线方法：杠杆（Addanki 等人，2022年）、ACE（Song 等人，2024年）、ABC3（Cha & Lee，2025年）、Causal-BALD（Jessonet al.，2021年）。对于决策场景，我们还纳入了 Sundin（Sundin 等人，2019年）、DEIG（Filstroff 等人，2024年）和 RFAN（Klein 等人，2025年）。

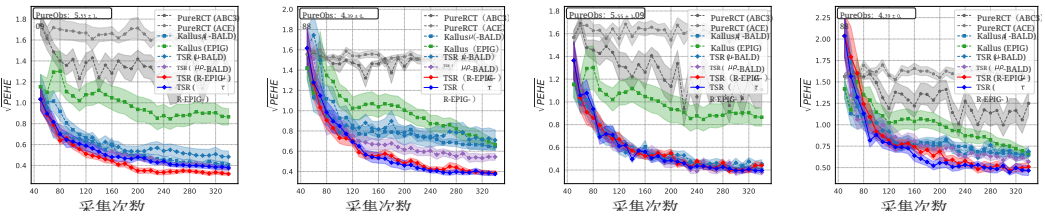


图4：比较两个试验估计器（CMGP 和 NSGP）在模拟数据集上的 √PEHE。（1、3）CMGP 和 NSGP；（2、4）存在严重协变量偏移时的 CMGP 和 NSGP。

推断架构与骨干模型。我们将提出的 R-Design（TSR）框架与两种基线架构进行比较：（1）PureRCT，这是一种标准方法，仅使用实验数据训练因果估计器，而忽略观测先验；（2）Kallus（Kallus 等人，2018年），这是一种数据融合方法，将观测数据和实验数据合并，并通过重要性加权校正混杂。所有框架均采用同一组稳健的贝叶斯学习器进行实例化。对于观测基线（阶段 1），我们主要使用 TabPFN-v2.5（Grinsztajn 等人，2025年），因为它在表格数据上表现出色，同时也评估 CausalPFN（Balazadeh 等人，2025年）。对于实验/残差学习阶段（阶段 2），我们采用能够提供有效联合后验分布的成熟贝叶斯估计器，包括 CMGP（Alaa 和 Van Der Schaar，2017）、NSGP（Alaa 和 Schaar，2018年）、BART（Hill，2011年）、BCF（Hahn 等人，2020年）和 CMDE（Jiang 等人，2023年）。

评估指标。我们使用两个主要标准评估性能。首先，报告 PEHE（式2），用于量化估计的处理效应相对于真实值的准确性。其次，对于决策任务，我们报告 APE（式3）以及观测总体上的平均遗憾（Gao 等人，2024b；Fernández-Loría 和 Provost，2022年）。所有结果均报告为 ±标准差，基于10次独立运行。

6.1 合成数据

我们使用两项模拟研究评估不同的采集函数：单变量设置和多变量设置。单变量研究包含 8 个数据集，涵盖基函数、偏差函数和CATE函数的简单与复杂组合，并评估CATE估计和策略学习任务上的性能。更多模拟结果见附录E，展示了R-Design在广泛设置中的鲁棒性与具有竞争力的性能。

结果。 图4 和7 总结了维度为 = 6 的协变量合成数据集上的主要结果，展示了我们提出的R-EPIG采集准则在TSR框架内于效应估计和策略学习任务上的强性能。

效应估计性能。如图4所示，我们提出的R-EPIG准则在整个采集轨迹上相较所有基线方法均持续实现显著的PEHE降低。在全部15种被评估方法中，基于Nemenyi事后检验的临界差异图显示，R-EPIG-μ 取得最佳平均排名3.00，R-EPIG-τ 以平均排名3.60位列第二。临界差异阈值（CD = 3.03，于 $\alpha = 0.05$）证实，两种R-EPIG 变体在统计上彼此无法区分，同时显著优于所有非TSR基线。值得注意的是，所有基于TSR的方法（排名1–7）均大幅优于基于Kallus的方法（排名8–10）和PureRCT方法（排名11–15），从而验证了我们利用观测数据进行目标分布估计的两阶段框架的有效性。成对Wilcoxon符号秩检验进一步证实，R-EPIG-μ以84.5%的胜率显著优于15个基线中的14个（$p < 0.05$），R-EPIG-τ 以80.5%的胜率优于13个基线。重要的是，两种R-EPIG 变体均未显著输给任何基线方法。

策略学习性能。对于处理策略优化，我们使用 APE 和平均遗憾（AR）评估指标来评估各方法。如图 5所示，R-EPIG-π是专为决策设计的策略感知采集函数，在这两项指标上均优于所有基线采集函数。这表明，直接以策略相关信息增益为目标，比将 CATE 估计准确度作为策略质量的代理指标更有优势。

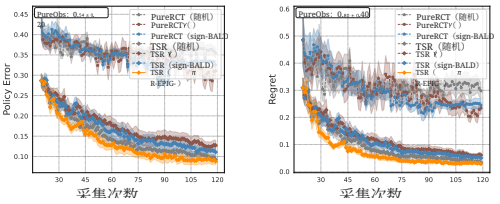


图5：所有方法在策略学习任务上的性能比较：(a) APE；(b) 平均遗憾。

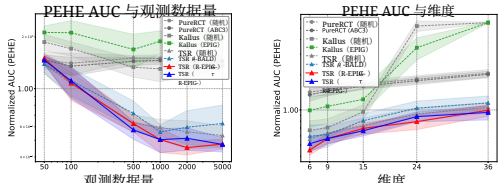


图6：不同先验规模和协变量维度下的曲线下面积。

可扩展性分析。图 6 考察了协变量维度和观测数据集大小变化时各方法的性能。当维度从 6 增加到 36时，TSR方法与基线方法之间的性能差距扩大，而 R-EPIG-μ 始终排名第一（平均排名 3。00），在所有维度下均保持这一表现。这种鲁棒性源于 R-EPIG 对目标分布的明确考量；在高维空间中，严重的分布偏移使这一点愈发重要。当观测数据量从 50 增加到 5000 个样本时，TSR方法随着观测数据增多而持续改进，同时各采集函数的相对排序保持稳定。

跨数据生成过程的鲁棒性。图3 评估了 R-EPIG 准则在以下27 种多样组合下的鲁棒性：基础效应函数（线性、多项式、正弦）、混杂偏差模式（均匀、中等、强）以及条件平均处理效应结构（常数、异质、交互）。在所有配置下，R-EPIG-μ和 R-EPIG-τ 始终名列前茅，表明其有效性并非局限于特定函数形式，而是能够推广到广泛的因果数据生成过程中。附录中提供了全面的数值结果、更多收敛图，以及详细的消融研究，用以考察对超参数、核函数选择和模型设定的敏感性。E。

表1：附加消融研究索引。

维度	评估的变体	Ref.
1. 主动学习配置		
批次大小	$n_b \in \{2, 3, 5, 10, 20, 30\}$	§18
池大小	$n_P \in \{500,1000,1500,2000,2500\}$	§16
2. 模型无关性（鲁棒性）		
阶段 1（先验）	TabPFN、CMGP、NSGP、CausalPFN	§15
阶段 2（试验）	CMGP、NSGP、BART、BCF、CMDE	§19
核类型	径向基函数、Matérn、有理二次	§14

更多消融研究。 表 1 总结了更多消融研究，考察我们的框架对各种超参数和模型选择的敏感性。这些研究涵盖主动学习配置（批次大小、池大小和温度），以及阶段 1（先验模型）和阶段 2（试验模型）的模型骨干网络变体。各项消融研究的详细实验设置、结果和分析见相应的附录章节。

6.2 半合成数据。

我们在两个公认的半合成基准上评估了该框架。第一个是 IHDP（Hill，2011年）（25 个协变量），通过从随机试验中移除一部分非随机选取的治疗组个体来引入选择偏倚。第二个是 ACTG-175（Hammer 等人，1996年）（12 个协变量），通过根据入组时的症状排除参与者来构建观察性队列。图8 报告了相较于采用随机获取的纯随机对照试验策略的相对性能提升。在 IHDP 上，采用 R-EPIG-μ 的 TSR 表现最佳，其次是 R-EPIG-τ。我们认为，这是因为 IHDP 的协变量维度更高，在这种情况下，以结果曲面为目标能够实现更稳健的不确定性量化。相比之下，ACTG-175 的协变量空间维度较低，因此 R-EPIG-τ 因直接关注处理效应异质性而排名第一，而 R-EPIG-μ

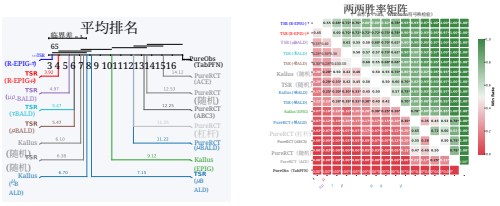


图 7：合成数据集上采集函数的比较：(a) 所有设置下的临界差异图，(b) 具有统计显著性的两两胜率矩阵

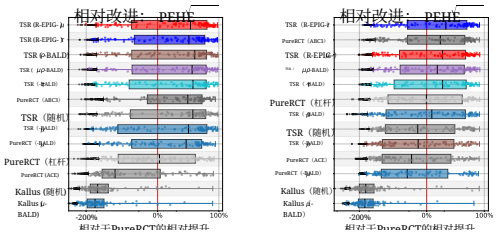


图 8：半合成数据集上采集函数的比较。IHDP 数据集（左）和 ACTG-175 数据集（右）上的相对性能改进。

排名第三，表明其在不同问题设定下仍具竞争力。更多结果见附录E。

7 相关工作

我们的框架处于因果推断中主动学习与自适应实验设计的交汇处。现有文献大致可根据学习器对干预机制的控制程度进行分类。一类研究聚焦于从固定的观察性队列中主动估计条件平均处理效应（Jesson等人，2021年；Gao 等人，2025年；Gao 与 Sejdinovic，2026年）。在这种设定下，处理已由未知策略分配，学习器则有选择地查询高成本结果，以改进条件平均处理效应模型。此处的核心挑战是一个结构性迁移问题：如何根据单个事实查询推断出成对的不可观测潜在结果。然而，学习器无法主动控制处理分配，因此直接探索反事实的能力受限。另一类截然不同的研究关注自适应实验设计（Addanki 等人，2022年；Zhang 等人，2025b）。在这类研究中，学习器能够分配处理。近期研究探索了旨在最小化平均处理效应方差的自适应随机化（Kato 等人，2024年）、有限样本池中的协变量平衡方法（Harshaw 等人，2024年），以及用于降低条件平均处理效应后验方差的贝叶斯实验设计（Cha & Lee，2025年；Klein 等人，2025年）。然而，这些方法通常采用白板设定，从头设计实验，同时忽略大量已有的观测数据。我们通过探讨一个混合型问题弥合这一差距：给定一个规模庞大但存在偏差的观测先验，应如何依次分配处理，才能以最高效率学习残差函数，从而消除先验偏差并恢复真实的条件平均处理效应？附录A中详细回顾了其他相关主题。

8 讨论与结论

我们提出了 R-Design，这一范式转变打通了大规模观察性研究与目标明确的实验设计之间的鸿沟。通过将问题形式化为主动残差学习，我们证明，有偏差的观测模型不应被弃用，而应重新用作基础先验。我们的理论分析严谨地确立了结构效率差距，证明估计平滑的残差对比所需的样本量，远少于从头重建复杂的结果机制。此外，R-EPIG 弥合了主动学习目标与下游因果任务之间的关键错位：通过严格聚焦目标总体，并针对具体目标（估计或策略）进行调整，它避免将预算浪费在无关的不确定性上。归根结底，R-Design 为资源受限条件下的因果推断提供了蓝图：它表明，我们不必通过大规模实验追求全局完美，而可以通过有策略地修正过去的偏差，实现更安全、更高效的决策。

未来工作。 虽然 R-Design 目前聚焦于具有单一先验的二元干预，但未来的迭代将扩展残差学习的范围。我们计划将该框架扩展到连续治疗方案（例如剂量），其中目标从估计标量对比转向学习复杂的残差剂量-反应曲面。

参考文献

Raghavendra Addanki、David Arbour、Tung Mai、Cameron Musco 与 Anup Rao。样本约束的处理效应估计。载于神经信息处理系统进展，第35卷，第5417–5430页，2022年。

Ahmed Alaa 和 Mihaela Schaar。异质性处理效应估计的局限：实用算法设计指南。载于国际机器学习会议，第 129–138 页。PMLR，2018年。

Ahmed M Alaa 和 Mihaela Van Der Schaar。使用多任务高斯过程对个体化处理效应进行贝叶斯推断。载于神经信息处理系统进展，第 30 卷，2017。

Susan Athey 和 Stefan Wager。使用因果森林估计处理效应：一项应用。观察性研究，5(2):37–51，2019年。

Vahid Balazadeh、Hamidreza Kamkari、Valentin Thomas、Benson Li、Junwei Ma、Jesse C. Cresswell 和 Rahul G. Krishnan。CausalPFN：通过上下文学习进行摊销因果效应估计。载于神经信息处理系统进展，第 38 卷，2025年。

Edwin V Bonilla、Kian Chai 和 Christopher Williams。多任务高斯过程预测。神经信息处理系统进展，20，2007。

Taehun Cha 和 Donghun Lee。ABC3：在随机实验中基于 Cohn 准则的主动贝叶斯因果推断。载于 *AAAI* 人工智能大会论文集，第 39 卷，第 26769–26777 页，2025年。

Bénédicte Colnet、Imke Mayer、Guanhua Chen、Awa Dieng、Ruohong Li、Gaël Varoquaux、Jean-Philippe Vert、Julie Josse 和 Shu Yang。结合随机试验与观察性研究的因果推断方法：综述。《统计科学》，39(1)：165–191，2024年。

Thomas Cook、Alan Mishler 和 Aaditya Ramdas。自适应实验中的半参数高效推断。载于 因果学习与推理，第 1033–1064 页。PMLR，2024年。

Alicia Curth 和 Mihaela Van der Schaar。异质性处理效应的非参数估计：从理论到学习算法。载于人工智能与统计国际会议，第 1810–1818 页。PMLR，2021年。

Irina Degtiar 和 Sherri Rose。可推广性与可迁移性综述。《统计学及其应用年度评论》，10(1):501–524，2023年。

Evangelos Dimitriou、Edwin Fong、Karla Diaz-Ordaz 和 Brieuc Lehmann。利用多任务高斯过程进行异质处理效应估计的数据融合。*arXiv* 预印本 *arXiv:2405.20957*，2024年。

Peng Ding。因果推断入门。CRC出版社，2024年。

Bradley Efron。使序贯实验达到平衡。*Biometrika*，58(3):403–417，1971年 .

Aristotelis Epanomeritakis 和 Davide Viviano。选择学习内容：结合实验与观察性证据时的实验设计。*arXiv* 预印本 *arXiv:2510.23434*，2025年。

Jake Fawkes、Michael O'Riordan、Athanasios Vlontzos、Oriol Corcoll 与 Ciarán Mark Gilligan-Lee。用实验数据验证观察性研究的困难性。载于人工智能与统计国际会议，《机器学习研究论文集》第258卷，第1819–1827页。PMLR，2025年5月3–5日，2025a。

Jake Fawkes、Lucile Ter-Minassian、Desi R. Ivanova、Uri Shalit 与 Christopher C. Holmes。合并是否值得？安全地评估因果数据集获取的信息增益。载于人工智能与统计国际会议，《机器学习研究论文集》第258卷，第1423–1431页。PMLR，2025年5月3–5日，2025b。

Carlos Fernández-Loría 和 Foster Provost。因果决策与因果效应估计并非一回事……以及这为何重要。*INFORMS*数据科学期刊，1(1):4–16，2022年。

Stefan Feuerriegel、Dennis Frauen、Valentyn Melnychuk、Jonas Schweisthal、Konstantin Hess、Alicia Curth、Stefan Bauer、Niki Kilbertus、Isaac S Kohane 和 Mihaela van der Schaar。用于预测治疗结局的因果机器学习。《自然·医学》，30(4):958–968，2024年。Louis Filstroff、Iiris Sundin、Petrus Mikkola、Aleksei Tiulpin、Juuso Kylmäoja 和 Samuel Kaski。面向贝叶斯决策的目标主动学习。《机器学习研究汇刊》，2024年。ISSN 2835-8856。

Jared C Foster、Jeremy MG Taylor 和 Stephen J Ruberg。基于随机临床试验数据的亚组识别。《医学统计学》，30(24):2867–2880，2011年。

Thomas R Frieden。卫生决策的证据——超越随机对照试验。《新英格兰医学杂志》，377(5):465–475，2017。

Chen Gao、Yu Zheng、Wenjie Wang、Fuli Feng、Xiangnan He 和 Yong Li。推荐系统中的因果推断：综述与未来方向。《*ACM*信息系统汇刊》，42(4):1–32，2024a。

Erdun Gao 和 Dino Sejdinovic。ActiveCQ：因果量的主动估计。发表于第十四届国际学习表征会议，2026年。

Erdun Gao、Howard Bondell、Wei Huang 和 Mingming Gong。用于在存在测量误差时估计连续处理效应的变分框架。发表于第十二届国际学习表征会议，2024b。

Erdun Gao、Jake Fawkes 和 Dino Sejdinovic。Causal-EPIG：一个用于条件平均处理效应估计的预测导向主动学习框架。*arXiv* 预印本 *arXiv:2509.21866*，2025年。

Mehrdad Ghadiri、David Arbour、Tung Mai、Cameron Musco 和 Anup B Rao。面向处理效应估计的有限总体回归调整与非渐近保证。神经信息处理系统进展，36:74180–74212，2023年。

GPy。GPy：Python 中的高斯过程框架。http://github.com/SheffieldML/GPy，自 2012 起。

Léo Grinsztajn、Klemens Flöge、Oscar Key、Felix Birkel、Philipp Jund、Brendan Roof、Benjamin Jäger、Dominik Safaric、Simone Alessi、Adrian Hayler 等。TabPFN-2.5：推进表格基础模型的最新技术水平。*arXiv* 预印本 *arXiv:2511.08667*，2025年。

Jinyong Hahn、Keisuke Hirano 和 Dean Karlan。使用倾向得分的自适应实验设计。《商业与经济统计学杂志》，29(1):96–108，2011年。

P Richard Hahn、Jared S Murray 和 Carlos M Carvalho。用于因果推断的贝叶斯回归树模型：正则化、混杂与异质性效应（附讨论）。《贝叶斯分析》，15(3):965–1056，2020年。

Scott M Hammer、David A Katzenstein、Michael D Hughes、Holly Gundacker、Robert T Schooley、Richard H Haubrich、W Keith Henry、Michael M Lederman、John P Phair、Manette Niu 等。比较核苷单药治疗与联合治疗对每立方毫米 CD4细胞计数为 200 至 500 的人类免疫缺陷病毒感染成人疗效的试验。《新英格兰医学杂志》，335(15)：1081–1090，1996年。Christopher Harshaw、Fredrik Sävje、Daniel A Spielman 和 Peng Zhang。采用 Gram–Schmidt随机游走设计在随机实验中平衡协变量。《美国统计学会杂志》，119(548)：2934–2946，2024年。Tobias Hatt、Jeroen Berrevoets、Alicia Curth、Stefan Feuerriegel 和 Mihaela van der Schaar。结合观测数据与随机化数据以估计异质性治疗效应。*arXiv* 预印本 *arXiv:2202.12891*，2022年。James J Heckman。经济学中的因果参数与政策分析：二十世纪回顾。《经济学季刊》，115(1)：45–97，2000年。

M.A. Hernan 和 J.M. Robins。因果推断：假如。Chapman & Hall/CRC 统计与应用概率专著。CRC出版社。ISBN 9781420076165。Jennifer L Hill。用于因果推断的贝叶斯非参数建模。计算与图形统计学期刊，20(1)：217–240，2011年。Noah Hollmann、Samuel Müller、Katharina Eggensperger 和 Frank Hutter。TabPFN：一种在一秒内解决小型表格分类问题的 Transformer。在国际学习表征会议 *2023*，2023年。Noah Hollmann、Samuel Müller、Lennart Purucker、Arjun Krishnakumar、Max Körfer、Shi Bin Hoo、Robin Tibor Schirrmeister 和 Frank Hutter。使用表格基础模型在小数据上实现准确预测。《自然》，2025年1月。doi: 10.1038/s41586-024-08328-6。URLhttps://www.nature.com/articles/s41586-024-08328-6。David Holzmüller、Viktor Zaverkin、Johannes Kästner 和 Ingo Steinwart。用于回归的深度批量主动学习框架与基准。机器学习研究期刊，24(164)：1–81，2023年。Neil Houlsby、Ferenc Huszár、Zoubin Ghahramani 和 Máté Lengyel。用于分类与偏好学习的贝叶斯主动学习。*arXiv* 预印本 *arXiv:1112.5745*，2011年。Jonas Hübotter、Lenart Treven、Yarden As 和 Andreas Krause。传导式主动学习：理论与应用。在神经信息处理系统进展，第 37 卷，第 124686–124755 页，2024年。Andrew Jesson、Panagiotis Tigas、Joost van Amersfoort、Andreas Kirsch、Uri Shalit 和 Yarin Gal。Causal-BALD：从观测数据中对结果进行深度贝叶斯主动学习，以推断处理效应。在神经信息处理系统进展，第 34 卷，第 30465–30478 页，2021年。

Ziyang Jiang、Zhuoran Hou、Yiling Liu、Yiman Ren、Keyu Li 和 David Carlson。使用多任务深度集成估计因果效应。发表于国际机器学习会议，第 15023–15040 页。PMLR，2023年。

Nathan Kallus、Aahlad Manas Puli 和 Uri Shalit。通过实验锚定消除隐藏混杂。神经信息处理系统进展，31，2018年。Motonobu Kanagawa、Philipp Hennig、Dino Sejdinovic 和 Bharath K Sriperumbudur。高斯过程与再生核：联系与等价性。*arXiv* 预印本 *arXiv:2506.17366*，2025年。Masahiro Kato、Takuya Ishihara、Junya Honda 和 Yusuke Narita。用于平均处理效应估计的高效自适应实验设计。*arXiv* 预印本 *arXiv:2002.05308*，2020年。Masahiro Kato、Kenichiro McAlinn 和 Shota Yasui。自适应双重稳健估计量及关于日志策略的悖论。神经信息处理系统进展，34：1351–1364，2021年。Masahiro Kato、Akihiro Oga、Wataru Komatsubara 和 Ryo Inokuchi。通过协变量选择进行处理效应估计的主动自适应实验设计。载于国际机器学习会议。PMLR，2024年。Edward H Kennedy。迈向异质性因果效应的最优双重稳健估计。电子统计学期刊，17(2)：3008–3049，2023年。Andreas Kirsch。用于回归的黑盒批量主动学习。《机器学习研究汇刊》，2023年。ISSN 2835-8856。专家认证。Andreas Kirsch、Sebastian Farquhar、Parmida Atighehchian、Andrew Jesson、Frédéric Branchaud-Charron 和 Yarin Gal。随机批次采集：深度主动学习的简单基线。《机器学习研究汇刊》，2023年。ISSN 2835-8856。专家认证。

Omer Noy Klein、Alihan Hüyük、Ron Shamir、Uri Shalit 和 Mihaela van der Schaar。迈向经监管机构确认的自适应临床试验：机器学习的机遇与解决方案。载于人工智能与统计国际会议，第4969–4977页。PMLR，2025年。Nikolay Krantsevich、Jingyu He 和 P Richard Hahn。用于估计异质性效应的随机树集成。载于人工智能与统计国际会议，第6120–6131页。PMLR，2023年。

Sören R Künzel、Jasjeet S Sekhon、Peter J Bickel 和 Bin Yu。使用机器学习估计异质性处理效应的元学习器。美国国家科学院院刊，116(10)：4156–4165，2019年。

Xinkun Nie 和 Stefan Wager。异质性处理效应的准预言机估计。*Biometrika*，108(2)：299–319，2021年。

Ezinne Nwankwo、Lauri Goldkind 和 Angela Zhou。针对复杂嵌入式结果进行因果推断的批次自适应标注。*arXiv* 预印本 *arXiv:2502.10605*，2025年。J Pearl。因果关系。剑桥大学出版社，2009年。Tian Qin、Tian-Zuo Wang 和 Zhi-Hua Zhou。预算约束下的异质性处理效应估计。载于国际机器学习会议，第8693–8702页。PMLR，2021年。

Evan TR Rosenman 和 Art B Owen。受观察性研究启发的实验设计。《因果推断杂志》，9(1):147–171，2021年。

Donald B Rubin。使用潜在结果进行因果推断：设计、建模与决策。《美国统计学会杂志》，100(469):322–331，2005年。

Dino Sejdinovic。使用核嵌入进行因果推断概述。载于计量经济学及相关主题的数据科学，《系统、决策与控制研究》，第 to appear 页。施普林格，2025年。

Uri Shalit、Fredrik D Johansson 和 David Sontag。个体治疗效应估计：泛化界与算法。载于国际机器学习会议，第 3076–3085 页。PMLR，2017。

Freddie Bickford Smith、Andreas Kirsch、Sebastian Farquhar、Yarin Gal、Adam Foster 和 To m Rainforth。面向预测的贝叶斯主动学习。载于人工智能与统计国际会议，第 7331–7348 页。PMLR，2023年。

Difan Song、Simon Mak 和 Jeff Wu。ACE：针对高成本实验的因果推断主动学习。载于第*29*届*ACM SIGKDD*知识发现与数据挖掘会议：因果推断与机器学习研讨会。ACM，2024年。Iiris Sundin、Peter Schulam、Eero Siivola、Aki Vehtari、Suchi Saria 和 Samuel Kaski。基于不平衡观测数据进行决策的主动学习。载于国际机器学习会议，第6046–6055页。PMLR，2019年。Max Tabord-Meehan。随机对照试验中用于自适应随机化的分层树。《经济研究评论》，90(5)：2646–2673，2023年。Christian Toth、Lars Lorch、Christian Knoll、Andreas Krause、Franz Pernkopf、Robert Peharz 和 Julius Von Kügelgen。主动贝叶斯因果推断。神经信息处理系统进展，35：16261–16275，2022年。Mark J van der Laan。自适应组序贯设计的构建与分析。2008年。Stefan Wager 和 Susan Athey。利用随机森林估计和推断异质性处理效应。《美国统计学会杂志》，113(523)：1228–1242，2018年。

Hechuan Wen、Tong Chen、Mingming Gong、Li Kheng Chai、Shazia Sadiq 和 Hongzhi Yin。通过主动学习增强处理效应估计：反事实覆盖视角。发表于国际机器学习会议，2025年。

Lili Wu 和 Shu Yang。结合实验研究与观察性研究的异质性处理效应集成 r-学习器。发表于因果学习与推理会议，第 904–926 页。PMLR，2022年。Shu Yang、Siyi Liu、Donglin Zeng 和 Xiaofei Wang。用于处理效应异质性和混杂函数的数据融合方法。伯努利，31(4)：2987–3012，2025年。Yun Yang 和 Surya T Tokdar。高维下的极小极大最优非参数回归。统计年刊，43(2)：652–674，2015年。Qiong Zhang、Yan Shuo Tan、Qinglong Tian 和 Pengfei Li。Tabpfn：一个模型统御一切？*arXiv* 预印本 *arXiv:2505.20003*，2025a。Zhiheng Zhang、Haoxiang Wang、Haoxuan Li 和 Zhouchen Lin。基于有限样本的主动处理效应估计。发表于第四十二届国际机器学习会议，2025b。Chuan Zhou、Yaxuan Li、Chunyuan Zheng、Haiteng Zhang、Min Zhang、Haoxuan Li 和 Mingming Gong。用于处理效应估计并处理未测量混杂的两阶段预训练-微调框架。发表于*ACM SIGKDD*知识发现与数据挖掘会议，第 2113–2123 页，2025年。

附录

目录

A 扩展相关工作 18B 实现细节 20B.1 自适应因果实验设计循环 .20B.2 模型架构 .21B.3 数据集 .23B.4 硬件.26**C 采集函数详情 26**

C.1 CATE 基线采集策略 .
26C.2 决策采集策略 .28C.3 R-EPIG：残差预期预测信息增益 .28C.4 计算复杂度 .
31

D 理论分析细节 31
D.1 将 R-Design 映射到 TAL 框架 .31D.2 可行不可约不确定性 .32D.3 定理2的证明 .
33D.4 引理1的证明 .35D.5 命题 2 的证明.37

E 更多实验结果 37E.1 合成数据集与消融实验 .38E.2 半合成数据集 .43**F 讨论：联合模型还是 TSR? 45**

A 扩展相关工作

本节将更详细地回顾与本文相关的文献，包括基于观测数据的条件平均处理效应估计、因果数据融合（观测数据与随机对照试验数据）以及自适应实验设计，并指出 R-Design 所要解决的具体差异和研究空白。

条件平均处理效应估计。估计条件平均处理效应的一种主流范式称为元学习器，它将问题分解为估计干扰成分（例如结果回归和倾向评分），然后根据伪结果构造最终的条件平均处理效应估计量（Curth & Van der Schaar, 2021）。代表性方法包括双重（T）学习器、单一（S）学习器和 X-学习器（Künzel 等人，2019）、领域学习器（Shalit 等人，2017），以及双重稳健方法，例如双重稳健学习器（Kennedy, 2023）和残差学习器（Nie & Wager, 2021）。除元学习器外，人们还提出了多种强大的替代方法，包括基于树的集成方法，如因果森林（Wager & Athey, 2018）和贝叶斯加性回归树（BART）（Hill，2011）；贝叶斯方法，如因果高斯过程（Alaa & Van Der Schaar, 2017; Alaa & Schaar, 2018）和贝叶斯因果森林（Hahn 等人，2020）；因果多任务深度集成（CMDE）（Jiang 等人，2023）；CausalPFN（Balazadeh 等人，2025）；以及基于核的方法（Sejdinovic, 2025）。虽然我们提出的数据采集策略与所有这些下游估计量兼容，并能使其受益，但我们的贡献与估计器设计相互独立：我们并非提出新的条件平均处理效应估计量，而是开发一种有原则的方法来收集信息量更高的数据，从而提高任何所选模型的样本效率。

条件平均处理效应估计的数据融合。观测数据与随机对照试验各自的优势互为补充，这推动了因果推断领域大量关于数据融合的研究。一条重要研究路线利用少量无混杂的随机对照试验数据，校正规模大得多的观测数据集中的隐藏混杂（Colnet 等人，2024）。一类方法会显式建模偏差或混杂函数，其假设范围从线性等强参数假设（Kallus 等人，2018）到更灵活的非参数估计量（Yang 等人，2025；Wu 和 Yang，2022）。另一种不同的范式基于表示学习，通常采用两阶段框架：先从大规模观测数据集中学习共享表示，再利用无混杂的随机对照试验数据加以优化或校准，由此产生了 CorNet（Hatt 等人，2022）和两阶段预训练-微调（TSPF）等方法（Zhou 等人，2025）。互补的贝叶斯视角则研究数据融合下验证的基本限度（Fawkes 等人，2025a），或采用多任务高斯过程来自适应地控制有偏数据源的影响（Dimitriou 等人，2024）。尽管这些研究为对固定异质数据集进行回顾性分析提供了强大工具，但我们的工作关注的是一个本质不同的前瞻性问题：如何利用现有观测数据集，主动引导未来实验数据的收集。

自适应因果实验设计。 大量研究致力于优化处理分配策略，以最大限度地降低预测误差。这些研究大致可分为四个方向。

(1) 自适应随机化： 一项主要研究方向是根据累积数据更新处理概率，以尽量降低估计量（通常为平均处理效应）的方差。这一领域建立在坚实的理论基础之上（van der Laan，2008年；Hahn 等人，2011年），近期研究提出了更精细的设计，利用干扰成分的在线估计（Kato 等人，2021年；Tabord-Meehan，2023年），专门设计的估计量（如 A2IPW）（Kato 等人，2020年），以及针对自适应数据收集定制的设计（Cook 等人，2024年）。
(2) 有限样本池实验设计： 另一项互补的研究方向考察如何为固定的有限个体样本池选择设计方案。该领域的研究涵盖了从对协变量平衡与鲁棒性之间权衡的基础分析（Efron，1971年）到提供有限样本保证的现代框架，例如杠杆得分采样（Addanki 等人，2022年；Ghadiri 等人，2023年）以及 Gram-Schmidt 游走（Harshaw 等人，2024年）。
(3) 条件平均处理效应的贝叶斯设计： 近期研究从贝叶斯实验设计的角度探索条件平均处理效应估计。这一研究方向中的方法将监管约束纳入考量（Klein 等人，2025年），或处理结构不确定性（Toth 等人，2022年）；另一些方法则开发基于高斯过程的采集函数，以最小化后验方差（Cha & Lee，2025年），或为网络场景提供保证（Zhang 等人，2025b）。然而，至关重要的是，这些方法通常采用白板设定，从头设计实验，而不利用先前的观测数据。
(4) 主动 *CATE* 估计： 最后，第四个研究方向聚焦于为条件平均处理效应估计主动获取结果；在此情形下，处理分配通常来自观测数据或已固定，但获取结果的成本较高（Nwankwo 等人，2025年）。其目标是有选择地查询队列中的结果，以改进 CATE 模型。一些现有方法采用标准主动学习启发式策略，例如基于多样性的核心集选择（Qin 等人，2021年；Wen 等人，2025年），或针对特定的因果误差分量（Sundin 等人，2019年）；但这些方法往往依赖间接代理指标，而非直接降低估计目标的不确定性。信息论准则与此目标更为契合：Causal-EPIG（Gao 等人，2025年）直接针对 CATE 估计目标的不确定性；Causal-EIG（Fawkes 等人，2025b）和 Causal-BALD（Jesson等人，2021年）则关注模型内部参数；ActiveCQ（Gao 与 Sejdinovic，2026年）利用高斯过程引导采集，将这一原则推广至更广泛的因果量。

观测数据辅助的实验设计。 尽管利用观测先验进行实验设计的混合范式已得到探索，但现有方法所依赖的假设与 R-Design 存在根本差异。例如，Rosenman &Owen (2021年) 利用观测数据指导实验分配，但其方法严格来说是静态的且参数化的。该方法采用预先定义的分层设置，仅使用观测数据构建层方差的稳健置信集，以优化单批次极小极大遗憾。类似地，Epanomeritakis & Viviano (2025年) 提出了一种选择实验以补充

算法 1 R-Design：基于残差的自适应因果实验设计 要求：观测数据 D_O，未标注候选池 $\mathrm{D}_P = \{\boldsymbol{x}_i\}_{i=1}^{n}$，预算 n_B，批次大小 n_b。确保：条件平均处理效应估计量 $\hat{\tau}(\cdot) = \hat{\tau}_o(\cdot) + \hat{\tau}_\delta(\cdot)$。// – 阶段 1：观测热启动 – 1：在 D_O 上训练观测学习器 $\hat{\mu}_o(\cdot, 0)$ 和 $\hat{\mu}_o(\cdot, 1)$。2：构建固定的观测对比：$\hat{\tau}_o(\boldsymbol{x}) \leftarrow \hat{\mu}_o(\boldsymbol{x}, 1) - \hat{\mu}_o(\boldsymbol{x}, 0)$。3：冻结 $\hat{\mu}_o(\cdot)$ 和 $\hat{\tau}_o(\cdot)$ 的参数，将其用作函数偏移。// – 阶段 2：自适应残差学习 – 4：初始化实验数据集 $\mathrm{D}_E \leftarrow \varnothing$。$\hat{}$ 5：初始化残差模型 $\boldsymbol{\delta}(\cdot)$（例如，用于向量 $\left[\delta(\cdot, 0)\ \delta(\cdot, 1)\right]^{\top}$ 的 MTGP）。6：当 $\left|\mathrm{D}_E\right| < n_B$ 和 $\mathrm{D}_P \neq \varnothing$ 时 // 1. 评估：按个体优化处理分配 7：初始化候选得分 $S \leftarrow \varnothing$ 和分配 $A \leftarrow \varnothing$。8：对于 $\boldsymbol{x} \in \mathrm{D}_P$ 中的每个候选项 9：计算两种处理组的效用：对于 $t \in \{0, 1\}$，计算 $u_t \leftarrow U(\boldsymbol{x}, t \mid \boldsymbol{\delta}, \mathrm{D}_E)$。10：选择最佳处理组：$t^* \leftarrow \arg\max_{t \in \{0,1\}} u_t$。11：记录最高得分：$S[\boldsymbol{x}] \leftarrow \max(u_0, u_1)$。12：记录分配：$A[\boldsymbol{x}] \leftarrow t^*$。13：结束循环 // 2. 选择：随机批次采集（Softmax 技巧）14：通过 Softmax 计算选择概率：$\boldsymbol{p} \leftarrow \mathrm{Softmax}(S/\mathrm{temp})$。15：根据 $\boldsymbol{p}$（不放回）从 D_P 中抽取大小为 n_b 的批次 $\mathrm{B}_x \subset$。16：构成最终查询批次：$\mathrm{B} \leftarrow \{(\boldsymbol{x}, A[\boldsymbol{x}]) \mid \boldsymbol{x} \in \mathrm{B}x\}$。17：查询实验结果：$\mathrm{Y}_{\mathcal{B}} \leftarrow \{y \mid (\boldsymbol{x}, t) \in \mathrm{B}\}$。18：$\mathrm{D}_P \leftarrow \mathrm{D}_P \setminus \mathrm{B}x$。// 3. 更新：残差计算与后验更新 19：计算残差元组：$\mathrm{D}_{\mathrm{new}} \leftarrow \{(\boldsymbol{x}, t, r) \mid (\boldsymbol{x}, t) \in \mathrm{B}\ \ y \in \mathrm{Y}_{\mathcal{B}}, r = y - \hat{\mu}_o(\boldsymbol{x}, t)\}$。20：更新 $\mathrm{D}_E \leftarrow \mathrm{D}_E \cup \mathrm{Dnew}$。$\hat{}$ 21：根据新数据 D_E 更新 $\boldsymbol{\delta}(\cdot)$。22：结束循环 $\hat{}$ 23：返回 $\hat{\tau}(\boldsymbol{x}) = \hat{\tau}_o(\boldsymbol{x}) + (\delta(\boldsymbol{x}, 1) - \delta(\boldsymbol{x}, 0))$ ·„„„

我们通过选择能够使估计效用最大化的处理组 $t^* \in \{0,1\}$，为每个候选 $\boldsymbol{x}$ 优化处理分配，确保将实验资源用于最具信息量的反事实。

批次选择策略。构造用于信息论采集函数的最优批次，通常需要耗费大量计算资源进行序贯贪心优化（每次选择后重新评估得分），以处理信息冗余（Holzmüller 等人，2023年）。为通过并行评分（每个周期仅计算一次得分）保持计算效率 S，我们采用通过 Softmax 技巧进行随机批次采集（Kirsch 等人，2023年）。我们不确定性地选取排名靠前的n_b 个候选项（这种方式有时会导致查询到由得分相近且较高的点构成的冗余簇），而是按照经过 Softmax 归一化的效用得分所确定的概率，从候选池中抽取批次 $\mathcal{B}$： $p(\boldsymbol{x}) \propto \exp(S[\boldsymbol{x}]/\tau)$。该策略引入必要的随机性，以近似实现批次多样性并有效覆盖输入空间，同时避免序贯重新评分带来的计算开销。

B.2 模型架构

从概念上讲，R-Design框架对于阶段 1 的观测学习器而言是模型无关的。然而，根据引理 1，此阶段的目标是将观测基线误差($n_O^{-\gamma_\mu}$ 降至最低。因此，我们将阶段 1 视为在大型观察性数据集 $\mathcal{D}_O$上解决标准的条件平均处理效应估计问题，并利用先进的估计量，最大限度地提高观测对比的准确性 $\hat{}\tau_o$。

阶段 1：观测学习器。在主要实验结果中，我们采用 TabPFN（v2.5）（Grinsztajn 等人，2025年），这是一种专为表格数据设计的最先进的先验数据拟合网络（PFN）（Zhang 等人，2025a）。我们采用 S-Learner 元策略，使用单个模型 $\hat{}\mu_o(\boldsymbol{x}, t)$ 在 D_O 上训练以预测结果，并将固定的观测对比定义为 $\hat{}\tau_o(\boldsymbol{x}) = \hat{}\mu_o(\boldsymbol{x}, 1) - \hat{}\mu_o(\boldsymbol{x}, 0)$。这一选择充分利用了 TabPFN 大规模预训练的优势，以捕捉复杂的协变量交互（Hollmann 等人，2023年；2025年）。为确保鲁棒性并进行全面评估，我们还对其他先进估计量进行基准测试，包括 CausalPFN（一种针对因果推断微调的 PFN 变体）（Balazadeh 等人，2025年）以及高斯过程估计器，例如 CMGP（Alaa 与 Van Der Schaar，2017年）和 NSGP（Alaa 和 Schaar，2018）。关键在于，无论采用何种架构，该模型都只在 D_O 上训练一次，并在随后的实验阶段作为固定偏移函数处理。

第二阶段：残差学习器（概率型条件平均处理效应估计器）。根据引理1，阶段2 的目标是估计残差过程 τ_δ。虽然形式上属于条件平均处理效应估计问题，但其结构截然不同：目标表示观测基线的偏差函数，其光滑度通常高于完整的结果曲面（Colnet 等人，2024年）。鉴于实验数据成本高且稀缺（$n_E \ll n_O$），我们需要一种数据效率高的模型。此外，为了驱动 R-EPIG，该模型必须是严格的概率型模型，能够提供联合后验分布（无论是通过闭式形式还是采样），以量化认知不确定性。为兼顾这些要求，我们采用 CMGP（Alaa 和 Van Der Schaar，2017）作为主要的残差学习器，因为它在小样本情形下表现稳定。为扩展至更大的预算，我们采用稀疏变分高斯过程（SVGP），并开展消融研究。我们还对其他概率型估计器进行基准测试，包括 NSGP（Alaa 和 Schaar，2018年）、BART（Hill，2011年）、BCF（Hahn 等人，2020年）和 CMDE（Jiang 等人，2023年）。

- **核架构：** 我们主要使用带自动相关性确定的径向基函数核。这一选择使模型能够学习各向异性长度尺度，从而有效捕捉残差偏置在不同协变量维度上的敏感度变化，正如引理所要求的。 1。我们还提供了使用不同核的消融研究。
- **超参数优化：** 每次批量采集后，都会通过最大似然（边缘似然最大化）优化核参数（长度尺度、信号方差）和似然噪声。
- **实现：** 所有高斯过程模型均使用`GPy` （GPy，自2012年起）[1]。

[1] https://github.com/SheffieldML/GPy

B.2.1 基线架构

为展示我们的 TSR 范式所作出的贡献，我们以两种基础架构基线进行基准比较：白板方法（标准实践）和参数化校正方法（在我们的实验中标记为 Kallus）Kallus 等人，2018年）（这是我们方法的结构前身）。

1. 白纸框架（直接估计）。 这代表实验设计中标准的"直接学习器"方法（Song 等人，2024年；Cha & Lee，2025年）。

• **结构：**它完全忽略观测数据 D_O（或仅将其用于协变量分布估计），遵循观察性相关因隐藏混杂而不可靠的原则。该学习器试图仅使用实验数据 D_E，从零开始重建完整的 CATE 函数曲面 $\tau(\boldsymbol{x})$。

• **实现：**我们采用与 R-Design 相同的 CMGP 主干模型，但直接在原始实验结果 $y \in \mathrm{D}_E$ 上对其进行训练，以预测 $\mu_e(\boldsymbol{x}, t)$。

• **局限性：**正如Kallus 等人（2018年）所指出的，该方法在理论上是一致的，但在统计上效率低下。它丢弃了大规模观测样本中包含的"全局信号"，迫使学习器耗费预算去重新学习$\mathcal{D}_O$ 与$\mathcal{D}_E$之间很可能共享的基本结果结构。

2. Kallus 校正框架（被动残差）。我们的 R-Design框架在结构上受到"用于 CATE 的递归划分"的启发，尤其借鉴了 Kallus 等人（2018年）提出的偏差校正策略。

• **结构：**与 R-Design 类似，该方法采用两阶段分解。
  – 阶段 *1*：从 D_O 中学习有偏的观测 CATE $\omega(\boldsymbol{x})$（等价于我们的 $\hat{}\tau_o(\boldsymbol{x})$）。– 阶段 *2*：使用实验数据 D_E 学习校正项 $\eta(\boldsymbol{x})$（等价于我们的 $\tau_\delta(\boldsymbol{x})$）。最终估计量构造为 $\hat{}\tau(\boldsymbol{x}) = \hat{}\omega(\boldsymbol{x}) + \hat{}\eta(\boldsymbol{x})$。

• **原始方法与我们的实现：**Kallus 等人的原始工作（2018年）假设混杂函数 $\eta(\boldsymbol{x}) = \theta^\top \boldsymbol{x}$ 采用参数化（通常为线性）形式，以确保在数据有限的情况下具有可识别性。为了提供与我们的 R-Design 主干模型容量相匹配的更强非线性基线，我们将第二阶段学习器从线性回归升级为标准 GP。

• **区别：** 尽管结构相似，但在我们的实验中，"Kallus 基线"代表了该架构的被动 （随机采样）或标准不确定性 （τ-BALD）版本。R-Design 将标准回归目标替换为残差感知主动学习目标（R-EPIG），明确针对校正项的认知不确定性，而非完整结果。此外，Kallus 专注于静态 CATE 估计。相比之下，我们的核心贡献是在两阶段框架内自适应地学习残差项，从而随着获取更多信息，动态优化 CATE 估计量。

B.2.2 CATE模型详情

本节介绍三类成熟的贝叶斯CATE估计量。我们详细说明了每种架构如何对处理效应建模，以及至关重要的一点：如何提取信息论采集函数所需的联合后验分布 $p(y,\ \tau(\boldsymbol{x}) \mid \boldsymbol{x},\ t,\ \mathrm{D})$ 或 $p(y,\ \delta(\boldsymbol{x}) \mid \boldsymbol{x},\ t,\ \mathrm{D})$。注意：此处以一般形式（对结果建模$y$ 和效应 τ）进行介绍，但在 *R-Design* 阶段 r中应用时，这些模型用于残差 r 和残差对比 τ_δ。

1. 贝叶斯树集成（BART 和 BCF）。[2] 基于树的方法可作为表格数据的稳健基线。我们使用 BART（Hahn 等人，2011年）及其因果扩展 BCF（Hahn 等人，2020年；Krantsevich 等人，2023年）。

• **条件平均处理效应估计：**BCF 显式地将条件期望参数化为 $\mathbb{E}[y \mid \boldsymbol{x}, t] = \mu(\boldsymbol{x}) + \tau(\boldsymbol{x})t$。此处，预后项 $\mu(\boldsymbol{x})$ 与处理效应 $\tau(\boldsymbol{x})$ 均被建模为

[2] https://github.com/socket778/XBCF

独立的树和先验。为缓解正则化诱导偏差（正则化诱导混杂），BCF 采用基于倾向得分的正交化，在经倾向得分调整的结果上拟合 $\mu(\boldsymbol{x})$。

• 通过高斯近似得到联合后验：BART/BCF 中的推断通过马尔可夫链蒙特卡洛进行，从而得到一组后验样本 $\{(\hat{y}(s), \hat{\tau}(s))\}_{s=1}^{S}$，而非闭式分布。为了计算采集函数所需的微分熵，我们采用因果主动学习中的标准近似方法（Jesson 等人，2021年；Kirsch，2023年；Gao 等人，2025年）。我们将马尔可夫链蒙特卡洛样本视为来自隐式多元高斯分布的抽样。对于候选点 $(\boldsymbol{x}, t)$，我们根据这些 S 样本，计算拼接向量 $[\hat{y}, \hat{\tau}]^{\top}$ 的经验均值向量和协方差矩阵。该高斯代理分布使我们能够高效地应用闭式互信息公式。

2. 基于高斯过程的估计器（CMGP 和 NSGP）。[3] 高斯过程是 R-Design（阶段 2）的主要概率基础，因为其具有解析上的易处理性（Kanagawa等人，2025年）。我们采用 CMGP（Alaa 和 Van Der Schaar，2017）和 NSGP（Alaa 和 Schaar，2018年）。

• **条件平均处理效应估计：** CMGP 将潜在结果视为一个多任务学习问题，并在条件均值向量 $\boldsymbol{\mu}(\boldsymbol{x}) = [\mu(\boldsymbol{x}, 0)\ \mu(\boldsymbol{x}, 1)]^{\top}$上设定联合高斯过程先验。它采用线性共区域化模型（LMC）核来捕捉处理臂之间的相关性。NSGP 通过使用非平稳核对联合空间 $\mathbf{X} \times \{0, 1\}$进行建模，以处理异质性。它直接在响应函数 $\mu(\boldsymbol{x}, t)$上设定高斯过程先验，其中核超参数取决于处理。在这两种情况下，条件平均处理效应均可由线性变换 $\tau(\boldsymbol{x}) = \mu(\boldsymbol{x}, 1) - \mu(\boldsymbol{x}, 0)$ 得出。

• **联合后验（闭式形式）：** 与树模型或集成模型不同，高斯过程可以提供精确的联合后验分布。由于高斯过程后验在经线性运算后仍保持闭合，观测结果 y（在 $\boldsymbol{x}, t$处）与未观测的目标估计量 $\tau(\boldsymbol{x})$ 的联合分布为多元高斯分布。我们根据高斯过程后验核以解析方式推导出均值向量和协方差矩阵，从而无需采用抽样近似，即可精确、高效地计算采集分数。

3. 因果多任务深度集成（CMDE）。[4] 对于高维任务，高斯过程的扩展性较差，我们可以采用 CMDE（Jiang 等人，2023年），它将深度集成的思路应用于因果推断。

• **条件平均处理效应估计：**CMDE 采用多头架构，包含三个组件：一个共享表征网络 $\phi(\boldsymbol{x})$（用于捕捉共同特征）以及两个处理特定头 $h_0(\phi(\boldsymbol{x}))$ 和 $h_1(\phi(\boldsymbol{x}))$。潜在结果建模为 $\hat{y}_t = h_t(\phi(\boldsymbol{x}))$。我们训练由$M$ 个独立初始化网络组成的集成，以量化认知不确定性。条件平均处理效应估计为两个头之间差值的平均值：$\hat{\tau}(\boldsymbol{x}) = \frac{1}{M}\sum_{m=1}^{M}(h_{(m)\,1}(\boldsymbol{x}) - h_{(m)\,0}(\boldsymbol{x}))$。
• **通过集成聚合构建联合后验：**深度集成无法提供解析后验。我们将集成成员视为来自潜在结果均值后验分布的近似样本。对于查询点 $\boldsymbol{x}$，每个集成成员 $m \in \{1, \ldots, M\}$ 都会输出均值向量的预测值 $\hat{\boldsymbol{\mu}}_{(m)}(\boldsymbol{x}) = [\hat{\mu}_{(m)}(\boldsymbol{x}, 0), \hat{\mu}_{(m)}(\boldsymbol{x}, 1)]^{\top}$。为构建联合后验，我们计算集成预测的经验矩。由此得到一个高斯近似：

$$p(\boldsymbol{\mu}(\boldsymbol{x}) \mid \mathcal{D}) \approx \mathcal{N}\left(\bar{\boldsymbol{\mu}}(\boldsymbol{x}), \hat{\boldsymbol{\Sigma}}(\boldsymbol{x})\right), \tag{16}$$

其中 $\bar{\boldsymbol{\mu}}(\boldsymbol{x}) = \frac{1}{M}\sum_{m=1}^{M}\hat{\boldsymbol{\mu}}^{(m)}(\boldsymbol{x})$ 是集成均值向量，$\hat{\boldsymbol{\Sigma}}(\boldsymbol{x})$ 是经验协方差矩阵。随后，可通过对该联合高斯分布进行线性变换，推导出 CATE $\tau(\boldsymbol{x})$ 的后验分布（以及它与任意候选结果 y 的协方差）（例如，$\tau = \mu_1 - \mu_0$）。

B.3 数据集

为全面评估 R-Design 在不同复杂度下的表现，我们采用多种验证策略。首先，我们从单变量合成场景入手，剖析特定

[3] https://github.com/vanderschaarlab/mlforhealthlabpub/tree/main/alg/ causal_multitask_gaussian_processes_个体处理效应 [4] https://github.com/jzy95310/ICK/tree/main/实验/causal_inference

函数形式（如非线性、混杂频率）对残差学习的影响。接着，我们转向多变量合成环境，评估其应对高维性和协变量偏移的鲁棒性。最后，我们在半合成数据集（IHDP、ACTG-175）上对性能进行基准测试。这些数据集将真实世界协变量结构与模拟结果相结合，并提供了包含真实因果估计量的严格测试场景。

B.3.1 单变量合成数据集

我们采用 Causal-ICM（Dimitriou 等人，2024年）中的模拟框架，构建一组由8 个场景组成的$\tau(\boldsymbol{x})$，系统地改变三个结构性组成部分：处理效应异质性 $\mu(\boldsymbol{x}, 0)$、基线结果 $\eta(\boldsymbol{x})$ 以及观测混杂机制 $\eta(\boldsymbol{x})$。

数据生成过程。 对于实验（RCT）源 $\mathrm{s} = e$，我们模拟试验参与模型：首先从 $\tilde{\mathbf{x}} \sim \mathrm{U}[-2, 2]$ 中抽取候选协变量，然后通过逻辑斯谛参与概率 $p(\mathrm{s} = e \mid \tilde{\mathbf{x}}) = \sigma(\beta_0 + \beta_1 \tilde{\mathbf{x}})$ 进行筛选。只有满足 $\mathrm{s} = e$ 的候选样本才能进入试验池。处理分配保持平衡：$\mathrm{t} \sim$ 伯努利(0.5)，结果遵循 $\mathrm{y} = \mu(\mathbf{x}, \mathrm{t}) + \epsilon$，其中 $\epsilon \sim \mathrm{N}(0, 1)$。对于观测源 $\mathrm{s} = o$，协变量从 $\mathbf{x} \sim \mathrm{U}[-2, 2]$ 中抽取，并采用混杂处理分配 $e(\boldsymbol{x}) = \sigma(\boldsymbol{x})$。观测结果存在结构性偏差：

$$\mu_o(\boldsymbol{x}, t) = \mu(\boldsymbol{x}, t) + (2t - 1)\eta(\boldsymbol{x}), \tag{17}$$

其中 $(2t - 1)\eta(\boldsymbol{x})$ 引入依赖于处理的偏移，以模拟未观测混杂。

分量函数。 我们为每个结构分量定义两个复杂度水平：

- **处理效应** $\tau(\boldsymbol{x}) = \mu(\boldsymbol{x}, 1) - \mu(\boldsymbol{x}, 0)$：– 简单：$\tau(\boldsymbol{x}) = 1 - \boldsymbol{x}$（线性）。– 复杂：$\tau(\boldsymbol{x}) = 1 + \boldsymbol{x} - \boldsymbol{x}^2$（非线性、非单调）。

- **基线结果** $\mu(\boldsymbol{x}, 0)$:
  - Simple: $\mu(\boldsymbol{x}, 0) = 0.8\boldsymbol{x}^3 - \boldsymbol{x}$ (Polynomial).
  - Complex: $\mu(\boldsymbol{x}, 0) = 2\sin(3\pi\boldsymbol{x}) - 1.5\exp(1.5(\boldsymbol{x} - 0.8))$ (High-frequency).
- **混杂偏倚** $\eta(\boldsymbol{x})$:
  - Simple: $\eta(\boldsymbol{x}) = 2\exp(-0.8(\boldsymbol{x} + 2))$ (Smooth decay).
  - Complex: $\eta(\boldsymbol{x}) = 5\exp\left(-\frac{(\boldsymbol{x}+3)^2}{12.5}\right)(1 + 0.2\cos(\boldsymbol{x}))$ (Oscillatory).

模拟情景。 我们评估这些分量的全因子组合（2^3），得到 8 种不同情景（表2）。

表 2：按组件复杂度定义的单变量模拟场景。

场景	混杂 $\eta(\boldsymbol{x})$	基线 $\mu(\boldsymbol{x}, 0)$	处理效应 $\tau(\boldsymbol{x})$
1	简单	简单	简单
2	简单	简单	复杂
3	简单	复杂	简单
4	简单	复杂	复杂
5	复杂	简单	简单
6	复杂	简单	复杂
7	复杂	复杂	简单
8	复杂	复杂	复杂

Sa本量。 随机对照试验候选池 D_P 包含 $n_P = 1000$ 个样本，而观察性 1
data集合D_O 包含 $n_O = 2000$ 个样本。结果在10个随机种子上取平均。

B.3.2 多变量合成数据集

我们将模拟扩展到多变量设置，以引入预后噪声、复杂混杂和协变量偏移等挑战。

特征结构。 我们定义了一个 d 维协变量空间 $\boldsymbol{x} \in \mathbb{R}^d$，并将特征划分为三组：

- **混杂因素C：** 同时影响 t 和 y 的特征（主要偏差来源）。
- **预后因素P：** 仅影响 y 的特征（方差来源）。
- **噪声特征 N：** 无关特征（鲁棒性测试）。

我们使用 $d = 6$ 进行主要实验（每组 2 个），并将规模扩大至 $d \in \{9,15,24,36\}$。

数据生成。 对于随机对照试验数据池，$\boldsymbol{x} \sim \mathrm{U}[-2,2]^d$ 采用平衡分配。对于观测数据，$\boldsymbol{x} \sim \mathrm{U}[-2,2]^d$，但分配通过 $\pi_o(\boldsymbol{x}) = \sigma(f_{\text{prop}}(\boldsymbol{x}_C) + 0.2\tanh(\sum \boldsymbol{x}_C))$ 引入混杂。结果生成方式为 $\mathrm{y} = \mu(\boldsymbol{x},\mathrm{t}) + \epsilon$(随机对照试验) 和 $\mathrm{y} = \mu_o(\boldsymbol{x},\mathrm{t}) + \epsilon$(观测数据)，其中 μ_o 包含偏差项。

基线结果 $\mu(\boldsymbol{x},0)$：

$$\mu(\boldsymbol{x},0) = f_\mu(\boldsymbol{x}_C, \boldsymbol{x}_P) + 0.5\sin(3\pi \boldsymbol{x}_{C,1})\cos(\pi \boldsymbol{x}_{C,2}). \tag{18}$$

这将随机多层感知机 f_μ 与高频分量结合起来，以挑战学习器 .

处理效应 $\tau(\boldsymbol{x})$：

$$\tau(\boldsymbol{x}) = 1 + 2\sin(2\pi \boldsymbol{x}_1) + 0.5\cos(\pi \boldsymbol{x}_2). \tag{19}$$

观察性偏差 $\eta(\boldsymbol{x})$：观察性均值偏移了 $\mu_o(\boldsymbol{x},t) = \mu(\boldsymbol{x},t) + (2t-1)\eta(\boldsymbol{x})$，其中：

$$\eta(\boldsymbol{x}) = 1.5 + (\boldsymbol{x}_{C,1} + \boldsymbol{x}_{C,2}) + 0.5\boldsymbol{x}_{C,1}^2. \tag{20}$$

这种偏差是平滑的（二次），使残差学习器能够有效地恢复 $\tau_\delta(\boldsymbol{x})$。

协变量偏移。 在偏移实验中，D_P 从 $\mathrm{U}[-4,2]^d$ 中抽取，而 D_O 保持为 $\mathrm{U}[-2,2]^d$，从而产生实验策略必须进行外推的区域。

方案。 $n_P = 1000, n_O = 2000$。主动学习：预热 50，批次大小为 10，迭代 30 次。对于策略评估（决策），我们采用预热 20，批次大小为 1，迭代 100 次。

B.3.3 半合成数据集

我们在两个源自真实世界研究的基准上验证 R-Design，这些基准既提供了逼真的协变量相关性，又保留了真实目标估计量。

IHDP 数据集。 基于婴儿健康与发展计划（Hill，2011年），经过预处理以引入选择偏倚（$n = 747$，$d = 25$）。

响应面。 潜在结果通过非线性指数模型生成 ：

$$\mu(\boldsymbol{x},0) = \exp\left((\boldsymbol{x} + 0.5)^\top \boldsymbol{\beta}\right), \tag{21}$$

$$\mu(\boldsymbol{x},1) = (\boldsymbol{x} + 0.5)^\top \boldsymbol{\beta} - \omega, \tag{22}$$

其中 $\boldsymbol{\beta}$ 是一个稀疏系数向量。

混杂机制。 对于观测数据划分，我们注入治疗依赖偏差 $\eta(\boldsymbol{x})$，使得 $\mu_o(\boldsymbol{x},\ t) = \mu(\boldsymbol{x},\ t) + (2t-1)\eta(\boldsymbol{x})$：

$$\eta(\boldsymbol{x}) = \gamma \cdot \left(1 + c_1 \boldsymbol{x}_1 + c_2 \boldsymbol{x}_2 + c_3 \boldsymbol{x}_1^2\right). \tag{23}$$

这里，$\boldsymbol{x}_1$，$\boldsymbol{x}_2$ 对应于标准化的“bw”和“b.head”。我们使用 $\gamma = 1.5$ 和随机化系数 c_i。

ACTG-175 数据集。 源自一项艾滋病临床试验（Hammer 等人，1996年）（$n = 813$， $d = 12$）。响应面。我们定义了一个复杂的结果曲面，将随机神经网络与可解释交互相结合：

$$\begin{aligned}\mu(\boldsymbol{x}, 0) = f_{\text{MLP}}(\boldsymbol{x}) + 6 + 0.3\boldsymbol{x}_{\texttt{wtkg}}^2 \\ - \sin(\boldsymbol{x}_{\texttt{age}})\,(\boldsymbol{x}_{\texttt{gender}} + 1) + 0.6\boldsymbol{x}_{\texttt{hemo}}\boldsymbol{x}_{\texttt{race}} - 0.2\boldsymbol{x}_{\texttt{z30}}.\end{aligned} \tag{24}$$

处理效应为 $\tau(\boldsymbol{x}) = 1.5\sin(\boldsymbol{x}_{\texttt{wtkg}})(\boldsymbol{x}_{\texttt{karnof_hi}} + 1) + 2\boldsymbol{x}_{\texttt{age}}$。

混杂与偏移。我们采用与 IHDP 相同的混杂偏差函数 $\eta(\boldsymbol{x})$。此外，我们通过对随机对照试验数据池中的老年患者进行欠采样（基于年龄中位数的概率拒绝采样），引入协变量偏移，以模拟资格偏差。

实验方案。数据按 4:1:2 的比例划分（观测数据训练集 / 观测数据测试集 / 随机对照试验数据池）。主动学习从 30 个热启动样本开始，进行 150 次单样本采集，并对 10 个随机种子的结果取平均。

B.4 硬件

计算资源。 所有实验均在一台符合以下规格的工作站上进行：

- **中央处理器：** Intel Core i9-13900KF（24 核，32 线程）
- **随机存取存储器：** 128GB
- **图形处理器：** NVIDIA GeForce RTX 4090（24 GB 显存）
- **操作系统：** Ubuntu 24.04 LTS
- **软件：** Python 3.10、PyTorch 2.3.1（CUDA 12.1）

C 采集函数细节

本节给出 R-Design 所用采集函数的详细推导与计算策略。我们聚焦两个主要目标：最小化条件平均处理效应函数的估计误差（**R-EPIG-Est**），以及最小化最优处理决策的不确定性（**R-EPIG-Policy**）。由于观测对比 $\hat{}\tau_o(\boldsymbol{x})$ 在实验阶段被视为固定偏移，完整条件平均处理效应 $\tau(\boldsymbol{x})$ 的不确定性完全由残差对比 $\tau_\delta(\boldsymbol{x})$ 的不确定性驱动。具体而言，由于 $\tau(\boldsymbol{x}) = \hat{}\tau_o(\boldsymbol{x}) + \tau_\delta(\boldsymbol{x})$，因此有：

$$\begin{aligned}\mathbb{V}[\tau(\boldsymbol{x}) \mid \mathcal{D}_E] &= \mathbb{V}[\tau_\delta(\boldsymbol{x}) \mid \mathcal{D}_E] \\ \mathrm{H}(\tau(\boldsymbol{x}) \mid \mathcal{D}_E) &= \mathrm{H}(\tau_\delta(\boldsymbol{x}) \mid \mathcal{D}_E) + \text{const.}\end{aligned} \tag{25}$$

因此，最大化关于残差 $\tau_\delta(\boldsymbol{x})$的信息增益，在数学上等价于最大化关于真实条件平均处理效应 $\tau(\boldsymbol{x})$ 的信息增益。

C.1 条件平均处理效应基线采集策略

我们将 R-Design 与多种采集策略进行比较，范围从随机采样到先进的信息论准则。为确保公平比较，所有基线均使用与我们的方法相同的底层高斯过程超参数（除非特定于不同的模型类别），并从候选池 D_P 中选择实验点 $(\boldsymbol{x}, t)$。

随机采样。 标准非自适应基准从候选池中均匀随机地选取候选对象：

$$\alpha_{\text{Random}}(\boldsymbol{x}, t) \coloneqq \mathcal{U}(0, 1). \tag{26}$$

这可作为量化主动适应收益的基本下界。

杠杆得分采样。一种来自随机线性代数的经典实验设计准则（Addanki 等人，2022年）。为将其适配到因果设定，我们定义核矩阵 $\boldsymbol{K}$ 于增广空间 $\mathcal{X} \times \{0,1\}$上。候选设计点 $(\boldsymbol{x}, t)$ 的杠杆得分与其在正则化帽子矩阵中的对角元成正比：

$$\alpha_{\text{Leverage}}(\boldsymbol{x}, t) \coloneqq \left[\boldsymbol{K}(\boldsymbol{K} + \sigma^2 \boldsymbol{I})^{-1}\boldsymbol{K}\right]_{(\boldsymbol{x},t),(\boldsymbol{x},t)}. \tag{27}$$

在高斯过程背景下，该分数与潜在结果$\mu(\boldsymbol{x},\ t)$的后验方差单调相关。最大化这一准则会选出对模型拟合影响最大的设计点，优先考虑当前估计最不稳定的区域（位于联合协变量-处理空间中）。

Causal-BALD 系列采集函数（Jesson等人，2021年）。

μ-BALD。一种启发式方法，会选择当前模型对潜在结果最不确定的点，而忽略处理组之间的相关性。对于预测$\mu(\boldsymbol{x},\ t)$的概率模型：

$$\alpha_{\mu\text{-BALD}}(\boldsymbol{x}, t) \coloneqq \max_{t' \in \{0,1\}} \mathbb{V}[\mu(\boldsymbol{x}, t') \mid \mathcal{D}_E]. \tag{28}$$

该方法侧重于降低全局结果不确定性，而非专门针对条件平均处理效应。

τ-BALD。通过最大化对比项 $\tau(\boldsymbol{x}) = \mu(\boldsymbol{x},\ 1) - \mu(\boldsymbol{x},\ 0)$ 的后验方差，直接针对处理效应的不确定性：

$$\alpha_{\tau\text{-BALD}}(\boldsymbol{x}) \coloneqq \mathbb{V}[\tau(\boldsymbol{x}) \mid \mathcal{D}_E]. \tag{29}$$

由于 τ 对查询所选择的特定处理臂（在模型空间中）保持不变，处理臂 t通常被随机分配，或被分配为使边际结果方差最大化。

ρ-BALD。基于Jesson等人（2021年）的推导，该准则优先选择模型对处理效应不确定（高 $\mathbb{V}[\tau]$）但对反事实结果有把握（低 $\mathbb{V}[\mu_{1-t}]$）的点。对于候选处理臂 t，得分与互信息下界成正比：

$$\alpha_{\rho\text{-BALD}}(\boldsymbol{x}, t) \coloneqq \frac{1}{2}\log\left(1 + \frac{\mathbb{V}[\tau(\boldsymbol{x}) \mid \mathcal{D}_E]}{\mathbb{V}[\mu(\boldsymbol{x}, 1-t) \mid \mathcal{D}_E]}\right). \tag{30}$$

请注意，与 τ-BALD不同，该得分取决于被查询的特定处理臂 t，并惩罚反事实结果 $1-t$ 高度不确定的候选点（因为这会使处理效应难以分离）。

μ-ρ BALD。该策略将 μ-BALD的全局探索与 ρ-BALD的针对性细化相结合。其定义为按信噪比对结果不确定性进行乘性缩放：

$$\alpha_{\mu\text{-}\rho}(\boldsymbol{x}, t) \coloneqq \mathbb{V}[\mu(\boldsymbol{x}, t) \mid \mathcal{D}_E] \cdot \frac{\mathbb{V}[\tau(\boldsymbol{x}) \mid \mathcal{D}_E]}{\mathbb{V}[\mu(\boldsymbol{x}, 1-t) \mid \mathcal{D}_E]}. \tag{31}$$

该准则会积极选择同时满足以下条件的点：预测结果不确定（$\mathbb{V}[\mu_t]$较高）、对条件平均处理效应具有信息量（$\mathbb{V}[\tau]$较高），并由有把握的反事实结果所锚定（$\mathbb{V}[\mu_{1-t}]$较低）。

ABC3（主动贝叶斯因果协变量条件化）。一种方差缩减准则，用于衡量获取单个观测后，目标总体中 CATE 不确定性的预期缩减量$\mathcal{D}_O$（Cha & Lee，2025年）。对于候选项（$\boldsymbol{x}_i$，t_i），其得分是 τ 在所有目标点 $\boldsymbol{x}_j \in \mathcal{D}_O$上的总期望方差缩减量：

$$\alpha_{\text{ABC3}}(\boldsymbol{x}_i, t_i) \coloneqq \sum_{\boldsymbol{x}_j \in \mathcal{D}_O} \frac{\text{Cov}[y(\boldsymbol{x}_i, t_i), \tau(\boldsymbol{x}_j) \mid \mathcal{D}_E]^2}{\mathbb{V}[y(\boldsymbol{x}_i, t_i) \mid \mathcal{D}_E]}. \tag{32}$$

选择总降幅较大的处理组 $t_i \in \{0,1\}$。

ACE（主动协方差利用）。优先选择跨处理臂协方差较强的点，着重关注观测一个处理臂就能为反事实处理臂提供大量信息的区域（Song 等人，2024年）：

$$\alpha_{\text{ACE}}(\boldsymbol{x}_i, t_i) \coloneqq \frac{|\text{Cov}[\mu(\boldsymbol{x}_i, 0), \mu(\boldsymbol{x}_i, 1) \mid \mathcal{D}_E]|}{\mathbb{V}[\mu(\boldsymbol{x}_i, t_i) \mid \mathcal{D}_E]}. \tag{33}$$

EPIG（期望预测信息增益）。一种信息论基线（Smith 等人，2023年，通过最大化候选观测 $(\boldsymbol{x}_i,\ t_i)$ 与观测目标集上的预测之间的互信息$\mathcal{D}_O$。它倾向于采集能够最大程度降低目标预测联合熵的点：

$$\alpha_{\text{EPIG}}(\boldsymbol{x}_i, t_i) \coloneqq \frac{1}{|\mathcal{D}_O|} \sum_{\boldsymbol{x}_j \in \mathcal{D}_O} \text{I}(y(\boldsymbol{x}_i, t_i); y(\boldsymbol{x}_j) \mid \mathcal{D}_E), \tag{34}$$

其中，$y(\boldsymbol{x}_j)$ 表示目标集内结果的预测分布。在高斯假设下，该互信息有闭式解：

$$\text{I}(y_i; y_j) = -\frac{1}{2} \log \left(1 - \frac{\text{Cov}[y(\boldsymbol{x}_i, t_i), y(\boldsymbol{x}_j)]^2}{\sigma_i^2 \sigma_j^2}\right). \tag{35}$$

C.2 决策采集策略

对于决策任务，我们关注以策略优化为目标的采集函数，而非全局条件平均处理效应估计准确度。这些方法旨在通过降低对处理效应符号 $\text{sign}(\tau(\boldsymbol{x}))$ 的不确定性来最小化 APE。

随机采样。与条件平均处理效应采集基线中的策略相同，从候选项 $(\boldsymbol{x},\ t)$ 中均匀随机选择。这是策略学习的非自适应基线。

Sign-BALD。针对二元处理决策的不确定性 $\pi(\boldsymbol{x}) = \mathbb{I}(\tau(\boldsymbol{x}) > 0)$。它通过后验分布最大化决策指示变量的方差（Jesson等人，2021年；Klein 等人，2025年）：

$$\alpha_{\text{Sign}}(\boldsymbol{x}) \coloneqq \mathbb{V}\left[\mathbb{I}(\tau(\boldsymbol{x}) > 0) \mid \mathcal{D}_E\right]. \tag{36}$$

实际上，对于生成 K 后验样本 $\{\tau^{(k)}(\boldsymbol{x})\}_{k=1}^{K}$ 的模型，我们会计算二元指示变量 $z^{(k)} = \mathbb{I}(\tau^{(k)}(\boldsymbol{x}) > 0)$ 的经验方差。当后验概率 $\mathbb{P}(\tau(\boldsymbol{x}) > 0) = 0.5$ 时，该分数达到最大值（0.25），这表示这些点恰好位于决策边界上，模型对最优处理的判断最为模糊。查询处理组 t 通常通过均匀随机方式选取。

Type-S 错误。选择在处理决策中最有可能产生“Type S”（Sign）错误的点（Sundin等人，2019年）。假设 CATE 的后验分布为高斯分布，$\tau(\boldsymbol{x}) \mid \mathcal{D}_E \sim \mathcal{N}(\hat{\mu}_\tau(\boldsymbol{x}),\ \hat{\sigma}_\tau^2(\boldsymbol{x}))$，采集分数定义为真实效应的符号与估计均值的符号不同的概率：

$$\alpha_{\text{Type-S}}(\boldsymbol{x}) \coloneqq \Phi\left(-\frac{|\hat{\mu}_\tau(\boldsymbol{x})|}{\hat{\sigma}_\tau(\boldsymbol{x})}\right), \tag{37}$$

其中 $\Phi(\cdot)$ 是标准正态累积分布函数。该准则优先选择信噪比 $|\hat{\mu}_\tau|/\hat{\sigma}_\tau$ 较低的点，这种情况要么发生在效应接近于零（决策边界）时，要么发生在不确定性 $\hat{\sigma}_\tau$ 极高时。

C.3 R-EPIG：残差期望预测信息增益

在正文中，我们提出了 R-EPIG 框架，将 EPIG（Smith 等人，2023年）和 Causal-EPIG（Gao 等人，2025年）中的信息论准则调整到这一新场景中。R-EPIG 采用自然迁移模式：候选集由实验池 $\mathcal{D}_P$构成，而目标集通常是观测总体$\mathcal{D}_O$，因为估计准确性在该总体中最为重要。

形式上，我们的目标是最大化目标总体上残差对比$\tau_\delta(\boldsymbol{x}^*)$后验熵的期望降低量。由于观测对比 $\hat{}\tau_o$ 是固定的，CATE τ 的不确定性完全由残差函数决定：$\mathrm{H}(\tau \mid \mathrm{D}_E) \equiv \mathrm{H}(\tau_\delta \mid \mathrm{D}_E)$。候选项 $(\boldsymbol{x}, t)$ 的采集分数定义为在目标集上取平均的期望信息增益：

$$\begin{aligned}\alpha_{\text{R-EPIG}}(\boldsymbol{x}, t) \coloneqq \mathbb{E}_{\boldsymbol{x}^* \sim \mathcal{D}_O}\, \mathbb{E}_{y \sim p(y \mid \boldsymbol{x}, t, \mathcal{D}_E)} \Big[& \mathrm{H}(\tau_\delta(\boldsymbol{x}^*) \mid \mathcal{D}_E) \\ & - \mathrm{H}(\tau_\delta(\boldsymbol{x}^*) \mid \mathcal{D}_E \cup \{(\boldsymbol{x}, t, y)\}) \Big] \qquad (38) \\ = \frac{1}{|\mathcal{D}_O|} \sum_{\boldsymbol{x}^* \in \mathcal{D}_O} & \mathrm{I}\left(y; \tau_\delta(\boldsymbol{x}^*) \mid \mathcal{D}_E\right). \qquad (39)\end{aligned}$$

我们提出了这一准则的两个变体，分别针对残差不确定性的不同方面 y.

C.3.1 R-EPIG-τ：针对残差效应不确定性

R-EPIG-τ 直接最大化候选实验结果 y（在 $\boldsymbol{x}$，t）与目标点处残差对比 $\tau_\delta(\boldsymbol{x}^*)$ 之间的互信息。在高斯过程后验下，观测结果与目标残差的联合分布为多元高斯分布。互信息有闭式表达式：

$$\mathrm{I}(y; \tau_\delta(\boldsymbol{x}^*)) = \frac{1}{2} \log \left(\frac{\sigma_y^2 \sigma_{\tau_\delta^*}^2}{\sigma_y^2 \sigma_{\tau_\delta^*}^2 - \mathrm{Cov}(y, \tau_\delta(\boldsymbol{x}^*))^2} \right) = \frac{1}{2} \log \left(\frac{1}{1 - \rho_{y, \tau_\delta^*}^2} \right). \qquad (40)$$

直观而言，该准则会选择结果与目标残差对比中的不确定性相关性最高的候选项（$\rho^2 \to 1$）。为高效实现这一准则，我们利用高斯过程核的性质展开残差协方差项。对于目标点 $\boldsymbol{x}^*$，残差对比定义为$\tau_\delta(\boldsymbol{x}^*) = \delta(\boldsymbol{x}^*, 1) - \delta(\boldsymbol{x}^*, 0)$。因此：

$$\sigma_{\tau_\delta^*}^2 = \mathbb{V}[\delta(\boldsymbol{x}^*, 1)] + \mathbb{V}[\delta(\boldsymbol{x}^*, 0)] - 2\mathrm{Cov}(\delta(\boldsymbol{x}^*, 1), \delta(\boldsymbol{x}^*, 0)). \qquad (41)$$

候选结果 y（在处理 t 下）与目标残差对比之间的交叉协方差为：

$$\mathrm{Cov}(y, \tau_\delta(\boldsymbol{x}^*)) = \mathrm{Cov}(\delta(\boldsymbol{x}, t), \delta(\boldsymbol{x}^*, 1)) - \mathrm{Cov}(\delta(\boldsymbol{x}, t), \delta(\boldsymbol{x}^*, 0)). \qquad (42)$$

关键的是，这些项可以直接从 CMGP 的后验协方差矩阵中提取（Alaa 和 Van Der Schaar，2017）或 NSGP 的后验协方差矩阵中提取（Alaa 和 Schaar，2018）。如果我们使用其他基础 CATE 估计器，例如 BART（Hill，2011年）、BCF（Hahn 等人，2020）和 CMDE（Jiang 等人，2023），我们也可以将联合后验近似为多元正态分布，并完全采用相同的计算方法来获得评分。

C.3.2 R-EPIG-μ：以联合残差结果为目标

或者，R-EPIG-μ 最大化关于残差潜在结果联合向量 $\boldsymbol{\delta}(\boldsymbol{x}^*) = [\delta(\boldsymbol{x}^*, 0) \;\; \delta(\boldsymbol{x}^*, 1)]^\top$ 的信息增益。此变体认识到，降低个体响应曲面的不确定性会隐含地约束效应差：

$$\alpha_{\text{R-EPIG-}\mu}(\boldsymbol{x}, t) \coloneqq \frac{1}{|\mathcal{D}_O|} \sum_{\boldsymbol{x}^* \in \mathcal{D}_O} \mathrm{I}\left(y; \boldsymbol{\delta}(\boldsymbol{x}^*) \mid \mathcal{D}_E\right). \qquad (43)$$

该向量目标的互信息通过舒尔补计算。设 $\boldsymbol{\Sigma}_{\boldsymbol{\delta}^*}$ 为2×2 目标残差的协方差矩阵。信息增益为：

$$\mathrm{I}(y; \boldsymbol{\delta}(\boldsymbol{x}^*)) = \frac{1}{2} \log \frac{\sigma_y^2}{\sigma_{y|\boldsymbol{\delta}^*}^2}, \quad \text{where } \sigma_{y|\boldsymbol{\delta}^*}^2 = \sigma_y^2 - \mathbf{c}^\top \boldsymbol{\Sigma}_{\boldsymbol{\delta}^*}^{-1} \mathbf{c}, \qquad (44)$$

以及 $\mathbf{c} = [\mathrm{Cov}(y, \delta(\boldsymbol{x}^*, 0)), \mathrm{Cov}(y, \delta(\boldsymbol{x}^*, 1))]^\top$。虽然由于每个目标都需要进行矩阵求逆，计算成本略高于 R-EPIG-τ，但这一变体能够捕捉联合输出空间中的相关性，而这些相关性可能会被差分算子抵消。

表 3：R-EPIG 变体比较。

方面	R-EPIG-τ	R-EPIG-μ
目标估计量	残差对比 τ_δ	残差向量 $\boldsymbol{\delta}$
信息度量	标量互信息	向量互信息
复杂度	$\mathrm{O}(n_P \cdot n_O)$	$\mathrm{O}(n_P \cdot n_O)$（常数级额外开销）
直觉	与条件平均处理效应的相关性最大	与结果的相关性最大

总结与实现策略

处理组选择。对于选定的候选对象 $\boldsymbol{x}$，我们计算两个处理组 $t \in \{0,1\}$ 的分数，并选择效用最高的处理组：$t^* = \arg\max_t \alpha_{\text{R-EPIG}}(\boldsymbol{x}, t)$。这使模型能够请求特定的反事实。例如，如果残差不确定性主要由 $\delta(\cdot,0)$ 决定，则观测对照组。

计算效率。基于高斯过程的实现有一项显著优势：所有方差和协方差项（σ^2、Cov）均可根据更新后的后验核 $\mathbf{K}_{\text{post}}$进行解析计算，无需蒙特卡洛抽样。在批处理设置中，我们采用附录 B.1中所述的“并行评分”，通过一次在图形处理器上进行的向量化操作，计算整个候选池 $\mathcal{D}_P$ 与目标集$\mathcal{D}_O$ 之间的协方差映射。

C.3.3 R-EPIG-π：针对策略不确定性

对于目标是优化处理策略 $\pi(\boldsymbol{x}) = \mathbb{I}(\tau(\boldsymbol{x}) > 0)$ 的决策任务，标准的方差缩减效率低下，因为它会过度关注效应符号已确定的区域。我们提出 R-EPIG-π，通过最大化目标点处最优策略决策的二元熵预期缩减来实现这一目标：

$$\alpha_{\text{R-EPIG-}\pi}(\boldsymbol{x}, t) \coloneqq \frac{1}{|\mathcal{D}_O|} \sum_{\boldsymbol{x}^* \in \mathcal{D}_O} \mathrm{I}\left(y; \pi(\boldsymbol{x}^*) \mid \mathcal{D}_E\right). \tag{45}$$

高斯—伯努利变量的推导。与估计情形不同，这里的互信息涉及连续的高斯结果 y 和离散的伯努利变量 π^*。我们将其表示为当前策略熵与后验熵期望之差：

$$\mathrm{I}(y; \pi) = \mathrm{H}(\pi(\boldsymbol{x}^*) \mid \mathcal{D}_E) - \mathbb{E}_{y \sim p(y \mid \boldsymbol{x}, t)}\Big[\mathrm{H}(\pi(\boldsymbol{x}^*) \mid y, \mathcal{D}_E)\Big]. \tag{46}$$

回顾一下，在 R-Design 中，CATE 分解为 $\tau(\boldsymbol{x}^*) = \hat{\tau}_o(\boldsymbol{x}^*) + \tau_\delta(\boldsymbol{x}^*)$。由于观测对比 τ_o ^ 是固定的，最优行动为处理（$t = 1$）的概率完全由残差高斯过程后验分布推导得出：

$$\mathbb{P}(\pi = 1) = \mathbb{P}(\tau_\delta(\boldsymbol{x}^*) > -\hat{\tau}_o(\boldsymbol{x}^*)) = \Phi\left(\frac{\mu_{\tau_\delta}(\boldsymbol{x}^*) + \hat{\tau}_o(\boldsymbol{x}^*)}{\sigma_{\tau_\delta}(\boldsymbol{x}^*)}\right) \coloneqq p_{\text{prior}}. \tag{47}$$

第一项就是二元熵 $\mathrm{H}_{\text{Bern}}(p_{\text{prior}})$。

通过求积计算期望后验熵。第二项需要对未观测到的实验结果 y 积分。由于高斯过程更新后的后验均值 $\mu_{\tau_\delta|y}$ 关于 y 呈线性，而后验方差 $\sigma^2_{\tau_\delta|y}$ 是确定的，因此条件概率 $p_{\text{post}}(y)$ 可表示为 y 的函数。我们使用高斯—埃尔米特求积来近似这个难以求解的积分。通过进行变量替换 $y = \mu_y + \sqrt{2}\sigma_y z$，可将期望转换为对 K 个标准正态节点的加权求和：

$$\mathbb{E}_y[\mathrm{H}_{\text{Bern}}(p_{\text{post}}(y))] \approx \frac{1}{\sqrt{\pi}} \sum_{k=1}^{K} w_k \cdot \mathrm{H}_{\text{Bern}}\left(\Phi\left(\frac{\mu_{\text{total}} + \frac{\mathrm{Cov}(y, \tau_\delta)}{\sigma_y}\sqrt{2} z_k}{\sigma_{\tau_\delta|y}}\right)\right), \tag{48}$$

其中 $\mu_{\text{total}} = \mu_{\tau_\delta}(\boldsymbol{x}^*) + \hat{\tau}_o(\boldsymbol{x}^*)$。在我们的实现中，我们使用 $K = 96$ 个节点，与 $\mathrm{O}(n^3)$ 的高斯过程推断相比，可提供很高的数值精度，同时计算开销可忽略不计。

表 4：R-EPIG 变体：条件平均处理效应估计与策略优化。

方面	R-EPIG-τ	R-EPIG-μ	R-EPIG-π
目标量	条件平均处理效应幅度	结果	策略
目标	最小化 PEHE	最小化 PEHE	最小化 APE
信息度量	高斯互信息	高斯向量互信息	伯努利互信息
计算	解析（闭式）	解析（闭式）	高斯-埃尔米特求积
关注区域	高认知不确定性（CATE）	高认知不确定性（结果）	决策边界

直觉与设计原理。 R-EPIG-π 的表现与其估计对应项截然不同。如表 4所示，它自然会关注决策边界（$\tau(\boldsymbol{x}) \approx 0$）。在这些区域中，$p_{\text{prior}} \approx 0.5$（最大熵），而来自 y 的少量信息增益就能显著消除符号歧义。相反，在具有强正向或负向效应的区域中，熵接近于零，因此采集受到抑制。

C.4 计算复杂度

R-EPIG 的一个关键设计选择是采用边缘均值形式（$\sum_{\boldsymbol{x}^*} \mathrm{I}(y;\cdot)$ 而非全局联合信息增益。

• **全局形式（难以处理）：** 计算 $\mathrm{I}(y;\boldsymbol{\delta}_{\text{global}})$ 需要针对整个目标集 (D_O) 计算更新后的协方差矩阵的行列式，并对每个候选查询执行，因此计算成本高达 $O(|\mathrm{D}_P| \cdot |\mathrm{D}_O|^3)$。

• **我们的求和方法（高效）：** 通过假设目标点之间的信息可加性（这是总信息的下界），我们将复杂度降低至 $O(|\mathrm{D}_P| \cdot |\mathrm{D}_O|)$。

关键在于，由于我们使用高斯过程，只需要候选池与目标集之间的预测交叉协方差。我们通过“并行评分”高效实现这一点：在单次批量图形处理器操作中计算 D_P 与 D_O 之间的完整核矩阵块 $\mathbf{K}_{PT}$，从而使 R-Design 能够扩展。

D 理论分析细节

本节将详细证明第5节中给出的理论结果，为 R-Design框架奠定严谨的理论基础。我们首先证明结构效率差距（引理1），正式说明在满足光滑度假设的条件下，学习残差对比 τ_δ 比从零开始的方法具有更高的样本效率。接下来，我们确立目标对齐（命题1），证明最小化贝叶斯 PEHE 风险与最小化残差对比的后验方差在数学上等价。最后，我们完整给出收敛分析，探讨 R-EPIG 采集策略（定理2）的收敛性，并将因果主动学习问题与通用的传导式主动学习（TAL）框架联系起来（Hübotter 等人，2024年）。

D.1 将R-Design映射到TAL框架

为运用TAL的理论工具，我们将基于残差的实验设计问题正式映射到高斯过程主动学习设定中。

1. 残差过程。我们将潜在残差函数 $\delta(\boldsymbol{x},\ t)$ 建模为定义在增广域 $\mathrm{X} = \mathrm{X} \times \{0,\ 1\}$ 上的高斯过程：

$$\delta(\tilde{\boldsymbol{x}}) \sim \mathcal{GP}(0, k_\delta(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{x}}')), \quad \text{where } \tilde{\boldsymbol{x}} \coloneqq (\boldsymbol{x}, t). \tag{49}$$

由于在实验阶段，观测基线 $\hat{}\mu_o(\boldsymbol{x},\ t)$ 被视为固定的函数偏移量，因此，观测结果 y 在数学上等价于观测残差过程的一次带噪实现：

$$y_i = \hat{\mu}_o(\boldsymbol{x}_i, t_i) + \delta(\boldsymbol{x}_i, t_i) + \epsilon_i \iff r_i \coloneqq y_i - \hat{\mu}_o(\boldsymbol{x}_i, t_i) = \delta(\boldsymbol{x}_i, t_i) + \epsilon_i. \tag{50}$$

因此，直接最小化残差对比 τ_δ 的后验方差，就能最小化 CATE τ的后验方差，因为 $\mathbb{V}[\tau(\boldsymbol{x}) \mid \mathrm{D}_E] = \mathbb{V}[\tau_\delta(\boldsymbol{x}) \mid \mathrm{D}_E]$（其中 $\tau_\delta(\boldsymbol{x}) \coloneqq \delta(\boldsymbol{x}, 1) - \delta(\boldsymbol{x}, 0)$）。

2. 空间与约束。 我们定义主动学习表述所涉及的集合：

• 增强池（D_P）： 所有潜在实验干预的集合。$\mathrm{D}_P = \{(\boldsymbol{x}_i,\ t) \mid i \in \{1 \ldots n\}\ t \in \{0,\ 1\}\}$。

• 有效实验设计（$\mathbb{S}_{\text{valid}}$）： 由于因果推断的基本问题，我们无法观测同一单位的两种潜在结果。当且仅当满足以下条件时，所选设计 $\mathrm{A} \subseteq \mathrm{D}_P$ 才有效：

$$(\boldsymbol{x}_i, 0) \in \mathcal{A} \implies (\boldsymbol{x}_i, 1) \notin \mathcal{A}, \quad \forall i \in \{1 \ldots n\}. \tag{51}$$

这对选择施加了分划拟阵约束。

• 增强目标集（$\boldsymbol{X}_{\text{tar}}$）： 我们希望预测以估计条件平均处理效应的残差集合。 $\boldsymbol{X}_{\text{tar}} = \{(\boldsymbol{x}^*,\ t) \mid \boldsymbol{x}^* \in \boldsymbol{X}_{\text{eval}} t \in \{0,\ 1\}\}$。

3. 弱次模性假设。为考虑观测分布与实验分布之间可能存在的分布偏移，我们采用一种宽松的次模性假设，这与近期的理论进展一致。

假设 2 (α-弱次模性)。 效用函数 $\psi(\mathcal{A}) = \mathrm{I}(\tau_\delta, \tilde{\boldsymbol{x}}_{\text{tar}}; \mathrm{y}_{\mathcal{A}})$ 是一个单调的，α-弱次模集合函数，其参数为 $\alpha \in (0, 1]$。也就是说，对于任意 $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{D}_P$ 以及 $v \in \mathcal{D}_P \setminus \mathcal{B}$：

$$\psi(\mathcal{A} \cup \{v\}) - \psi(\mathcal{A}) \geq \alpha \left(\psi(\mathcal{B} \cup \{v\}) - \psi(\mathcal{B})\right). \tag{52}$$

论证： 在 Pool $\subseteq$ Target 的归纳式设置中，互信息具有严格次模性（$\alpha = 1$），但在直推式设置中，由于分布偏移及潜在信息协同作用，这一性质可能放宽为弱次模性（$\alpha < 1$）（Hübotter 等人， 2024年，附录 C.4）。假设$\alpha > 0$ 是一种标准的放宽形式，同时保留了贪心策略的常数因子近似保证。

D.2 可行的不可约不确定性

自适应因果实验设计面临一个独特挑战："不可约"不确定性并非样本空间的静态属性，而是取决于处理分配的选择。由于我们无法观测同一个体的两种潜在结果（Gao 等人，2025年），我们最终获取的信息取决于选择揭示哪种潜在结果。与标准 TAL 不同，在标准 TAL 中 $\eta_{\mathcal{S}}^2$ 是固定的（观测整个样本空间 $\mathcal{S}$），而在这里，我们只能观测满足因果约束的子集 $\tilde{\mathcal{D}} \subset \mathcal{D}_P$。为严格定义收敛极限，我们考虑：如果对整个样本池进行最优实验，所能实现的最佳方差缩减。

定义 1 (极大有效设计)。 设 $\mathcal{D}_P$ 为大小为 $2n$ 的增强池（包含所有潜在结果）。若设计 $\mathcal{D} \subset \mathcal{D}_P$ 对池中的每个单位恰好包含一个结果，则称其为极大有效设计：

$$\begin{aligned} \mathcal{D} \in \mathbb{S}_{max} &\iff |\mathcal{D}| = n \\ \forall i \in \{1 \ldots n\}, |\mathcal{D} \cap \{(\boldsymbol{x}_i, 0), (\boldsymbol{x}_i, 1)\}| &= 1. \end{aligned} \tag{53}$$

现在，我们将可行不可约不确定性 η_{feas}^2 定义为所有极大有效设计中可达到的最小方差。它是任何实验策略的预言机下界。

定义 2 (可行极限)。 目标 $\boldsymbol{x}^*$ 的可行不可约不确定性定义为：

$$\eta_{feas}^2(\boldsymbol{x}^*) \coloneqq \min_{\mathcal{D} \in \mathbb{S}_{max}} \mathbb{V}[\tau_\delta(\boldsymbol{x}^*) \mid \mathcal{D}]. \tag{54}$$

备注： 在标准TAL框架（Hübotter 等人，2024年）中，不可约不确定性对应于观测完整的样本空间。在我们的因果设定中，η_{feas}^2 表示最优可行极限，即理想实验者将处理最优地分配给整个总体后剩余的不确定性。我们的收敛分析（定理 2）证明，greedyR-EPIG 策略能够高效地逼近这一预言机极限。

D.3 证明定理 2

我们给出了一份详细证明，确立了 R-EPIG采集策略下后验方差的一致收敛。我们的证明策略将残差核的谱衰减 （引理1）与信息论主动学习的贪心优化保证相衔接（Hübotter 等人，2024年）。

预备知识。令 $\gamma_n(\tau_\delta)$ 表示经过 n 次观测后残差对比函数的最大信息容量。令 $\Gamma_n \coloneqq \max_{(\boldsymbol{x},t)} \mathrm{I}(\tau_\delta; \mathrm{y}_{\boldsymbol{x},\ t} \mid \mathrm{D}_{n-1})$ 表示在步长 n可获得的最大边际信息增益。我们旨在界定获得 n_B 个数据点后，最大后验方差 $\sup_{\boldsymbol{x}^* \in \mathcal{X}} \mathbb{V}[\tau_\delta(\boldsymbol{x}^*) \mid \mathrm{D}_{n_B}]$。

步骤 1：通过近似马尔可夫边界 (AMB) 化简问题。核心挑战在于将观测点上的信息增益转化为对任意未观测目标 $\boldsymbol{x}^* \in \mathrm{X}$ 的不确定性降低。我们引入 AMB 的概念。由于 τ_δ 属于一个特征值快速衰减的再生核希尔伯特空间，其有效维数较小。形式上，对于任意目标 $\boldsymbol{x}^*$ 和容差 $\epsilon > 0$，存在一个有限子集 $\mathrm{B}_\epsilon(\boldsymbol{x}^*) \subset \mathrm{D}_P$ （即 AMB），其大小为 $|\mathrm{B}_\epsilon| =: k_\epsilon$，使得在给定 B_ϵ 后，关于 $\boldsymbol{x}^*$ 的剩余信息可以忽略不计：

$$\mathrm{I}(\tau_\delta(\boldsymbol{x}^*); \mathrm{y}_{\tilde{\mathcal{D}}_P} \mid \mathrm{y}_{\mathcal{B}_\epsilon}) \leq \epsilon. \tag{55}$$

利用高斯过程（GP）中熵与方差之间的关系（Hübotter 等人，2024年，引理 C.17），$\boldsymbol{x}^*$ 处的后验方差由 AMB 中剩余的信息以及不可约近似误差所界定：

$$\mathbb{V}[\tau_\delta(\boldsymbol{x}^*) \mid \mathcal{D}] \leq 2\sigma^2 \cdot \mathrm{I}(\tau_\delta(\boldsymbol{x}^*); \mathrm{y}_{\mathcal{B}_\epsilon} \mid \mathcal{D}) + \eta^2_{\text{feas}}(\boldsymbol{x}^*) + C_1\epsilon, \tag{56}$$

其中，D 是任意历史信息，η^2_{feas} 是可行不可约不确定性。

步骤 2：R-EPIG 的优化保证。现在我们对以下项进行界定：$\mathrm{I}(\tau_\delta(\boldsymbol{x}^*);\ \mathrm{y}_{\mathcal{B}_\epsilon} \mid \mathcal{D}_{n_B})$。R-EPIG 算法通过贪心策略最大化信息增益。在 α-弱次模性（假设2）的前提下，贪心选择 $\mathcal{D}_{n_B}$ 与最优集合相比具有竞争力。具体而言，贪心策略的信息保留量受当前边际增益 Γ_{n_B} 按 AMB 的大小缩放后所界定。根据Hübotter 等人(2024年，引理 C.18）：

$$\mathrm{I}(\tau_\delta(\boldsymbol{x}^*); \mathrm{y}_{\mathcal{B}_\epsilon} \mid \mathcal{D}_{n_B}) \leq \frac{k_\epsilon}{\alpha}\Gamma_{n_B}. \tag{57}$$

将其代入式56，可得一个与边际增益相关的逐点界：

$$\mathbb{V}[\tau_\delta(\boldsymbol{x}^*) \mid \mathcal{D}_{n_B}] \leq \eta^2_{\text{feas}}(\boldsymbol{x}^*) + \frac{2\sigma^2 k_\epsilon}{\alpha}\Gamma_{n_B} + C_1\epsilon. \tag{58}$$

步长 3：收敛速率分析。我们利用边际增益之和受总容量上界限制这一性质：$\sum_{t=1}^{n_B} \Gamma_t \leq \gamma_{n_B}(\tau_\delta)$。由于边际增益非递增（这是次模性的结果），步长 n_B 处的瞬时增益上界为平均增益：$\Gamma_{n_B} \leq \gamma_{n_B}(\tau_\delta)/n_B$。

为推导均匀界，我们平衡近似误差 ϵ 与估计误差（其随 k_ϵ变化）。对于具有多项式谱衰减的核（这与引理 1中的霍尔德光滑性假设一致），AMB 的规模为 $k_\epsilon \approx \mathcal{O}(\epsilon^{-p})$，其中某个特征维数为 p。遵循标准的平衡论证（Hübotter 等人，2024年，定理3.3），我们选择 $\epsilon \asymp n_B^{-1/(1+p)}$。为简化说明，假设标准衰减情形下 $\epsilon \asymp n_B^{-1/2}$ 是最优的，我们得到：

$$\sup_{\boldsymbol{x}^*} \mathbb{V}[\tau_\delta(\boldsymbol{x}^*) \mid \mathcal{D}_{n_B}] \leq \eta^2_{\text{feas}}(\boldsymbol{x}^*) + \underbrace{C_2 \frac{\sqrt{n_B} \cdot \gamma_{n_B}(\tau_\delta)}{n_B}}_{\text{Estimation}} + \underbrace{C_1 n_B^{-1/2}}_{\text{Approximation}}. \tag{59}$$

简化估计项可得最终速率：

$$\sup_{\boldsymbol{x}^*} \mathbb{V}[\tau_\delta(\boldsymbol{x}^*) \mid \mathcal{D}_{n_B}] \leq \eta^2_{\text{feas}}(\boldsymbol{x}^*) + \tilde{\mathcal{O}}\left(\frac{\gamma_{n_B}(\tau_\delta)}{\sqrt{n_B}}\right). \tag{60}$$

至关重要的是，正如引理 1所确立的，残余容量 $\gamma_{n_B}(\tau_\delta)$ 的增长显著慢于完整结果容量 $\gamma_{n_B}(\tau)$，这是由于更高的光滑性（$\alpha_\delta > \alpha_{\mu_e}$），从而形式化地证明了R-Design框架的效率增益。

□

D.3.1 命题1的证明。1

在本小节中，我们严格证明：最小化期望贝叶斯PEHE在数学上等价于最小化残差对比$\tau_\delta(\boldsymbol{x})$的期望积分后验方差。

设定与定义。令 $\mathrm{H}_k = \mathrm{D}_O \cup \mathrm{D}_E^{(k)}$ 表示截至步长 k 收集到的历史信息。考虑一个候选实验设计 $(\boldsymbol{x},\ t)$。令 y 为预期结果，其服从后验预测分布 $p(y \mid \boldsymbol{x},\ t,\ \mathrm{H}_k)$。增广后的历史信息为 $\mathrm{H}_{k+1} = \mathrm{H}_k \cup \{(\boldsymbol{x},\ t,\ y)\}$。贝叶斯 CATE 估计量定义为处理效应的后验均值：

$$\hat{\tau}_{k+1}(\boldsymbol{x}^*) \coloneqq \mathbb{E}[\tau(\boldsymbol{x}^*) \mid \mathcal{H}_{k+1}]. \tag{61}$$

目标是最小化下一步的预期贝叶斯 PEHE 风险。该期望涵盖两种不确定性来源：未观测到的未来结果 y（预测不确定性）以及真实因果效应 τ（认知不确定性）：

$$\mathcal{J}(\boldsymbol{x},t) = \mathbb{E}_{y\sim p(y|\boldsymbol{x},t,\mathcal{H}_k)}\left[\mathbb{E}_{\tau\sim p(\tau|\mathcal{H}_{k+1})}\left[\int_{\mathcal{X}} \left(\hat{\tau}_{k+1}(\boldsymbol{x}^*) - \tau(\boldsymbol{x}^*)\right)^2 p_{\text{tar}}(\boldsymbol{x}^*)d\boldsymbol{x}^*\right]\right]. \tag{62}$$

证明。证明分为两步：首先，我们推导平方误差风险与后验方差之间的关系；其次，我们应用 R-Design 的结构分解，以分离残差方差。

步骤 1：建立 PEHE 与 τ 的后验方差之间的关系。 假设满足标准正则条件（有界矩假设与 Cha & Lee (2025 年，定理 4.1 ））， 我们应用 Fubini 定理，交换对 $\mathcal{X}$ 的积分与期望。对于给定评估点 $\boldsymbol{x}^*$，在增强后的历史信息 $\mathcal{H}_{k+1}$条件下，我们考察内层风险项。为简洁起见，记 $\hat{}\tau \coloneqq \hat{}\tau_{k+1}(\boldsymbol{x}^*)$ 和 $\tau \coloneqq \tau(\boldsymbol{x}^*)$。逐点风险为：

$$\mathrm{Risk}(\boldsymbol{x}^* \mid \mathcal{H}_{k+1}) = \mathbb{E}_{\tau|\mathcal{H}_{k+1}}\left[(\hat{\tau} - \tau)^2\right]. \tag{63}$$

由于估计量 $\hat{}\tau$ 是后验均值，因此在给定 H_{k+1}的条件下，它是一个确定值。我们展开二次项：

$$\begin{aligned}\mathbb{E}_\tau\left[(\hat{\tau} - \tau)^2\right] &= \mathbb{E}_\tau\left[\hat{\tau}^2 - 2\hat{\tau}\tau + \tau^2\right] \\ &= \hat{\tau}^2 - 2\hat{\tau}\underbrace{\mathbb{E}_\tau[\tau]}_{=\hat{\tau}} + \mathbb{E}_\tau[\tau^2] \\ &= \mathbb{E}_\tau[\tau^2] - (\mathbb{E}_\tau[\tau])^2.\end{aligned} \tag{64}$$

根据定义， $\mathbb{E}[\tau^2] - (\mathbb{E}[\tau])^2$ 是 τ的后验方差。因此，我们明确得到：

$$\mathrm{Risk}(\boldsymbol{x}^* \mid \mathcal{H}_{k+1}) = \mathbb{V}[\tau(\boldsymbol{x}^*) \mid \mathcal{H}_{k+1}]. \tag{65}$$

将其代回全局目标 J，优化问题便简化为最小化 τ的期望积分后验方差：

$$\arg\min_{(\boldsymbol{x},t)} \mathbb{E}_y\left[\int_{\mathcal{X}} \mathbb{V}[\tau(\boldsymbol{x}^*) \mid \mathcal{H}_{k+1}]\, p_{\text{tar}}(\boldsymbol{x}^*)d\boldsymbol{x}^*\right]. \tag{66}$$

步长 2：观测偏移论证。回顾基本的结构分解：$\tau(\boldsymbol{x}^*) = \hat{}\tau_o(\boldsymbol{x}^*) + \tau_\delta(\boldsymbol{x}^*)$。我们分析在给定历史信息 H_{k+1} 条件下的方差项。请注意，观察性数据集是历史信息的子集 $(\mathrm{D}_O \subset \mathrm{H}_k \subset \mathrm{H}_{k+1})$。由于观测对比学习器仅使用 D_O 进行训练，且其参数此后保持固定，因此在给定 H_{k+1} 的条件下，项 $\hat{}\tau_o(\boldsymbol{x}^*)$ 是一个确定的常数（它不依赖于新的结果 y）。应用方差的平移不变性 $(\mathbb{V}[X + c] = \mathbb{V}[X])$：

$$\begin{aligned}\mathbb{V}[\tau(\boldsymbol{x}^*) \mid \mathcal{H}_{k+1}] &= \mathbb{V}[\hat{\tau}_o(\boldsymbol{x}^*) + \tau_\delta(\boldsymbol{x}^*) \mid \mathcal{H}_{k+1}] \\ &= \mathbb{V}[\tau_\delta(\boldsymbol{x}^*) \mid \mathcal{H}_{k+1}].\end{aligned} \tag{67}$$

这一步证明，总CATE τ的后验不确定性完全由残差对比 τ_δ的后验不确定性决定。

结论。将步长 2 的结果代入公式66，最终目标为：

$$\arg\min_{(\boldsymbol{x},t)} \mathbb{E}_y\left[\int_{\mathcal{X}} \mathbb{V}[\tau_\delta(\boldsymbol{x}^*) \mid \mathcal{D}_O, \mathcal{D}_E^{(k)} \cup \{(\boldsymbol{x},t,y)\}]\, p_{\text{tar}}(\boldsymbol{x}^*) d\boldsymbol{x}^*\right]. \tag{68}$$

这证实，最小化预期条件平均处理效应估计误差，在数学上等同于最小化残差对比 $\tau_\delta(\boldsymbol{x}^*)$ 的预期积分后验方差。 □

D.3.2 理论依据：信息增益作为代理指标

虽然命题1 确立了积分方差缩减（IVR）是最小化 PEHE 的精确理论目标，但直接优化这一目标会带来算法上的挑战。IVR 目标对应于 A-最优性（最小化后验协方差的迹，Tr(Σ)），而它通常不具备次模性，这意味着贪心策略缺乏理论保证。我们提出的 R-EPIG 准则则旨在最大化目标总体中有关残差对比的互信息 $\mathrm{I}(y;\tau_\delta, \mathcal{X}_{\text{tar}})$。对于高斯过程，最大化互信息等价于最小化预期后验熵，而这对应于 D-最优性（最小化后验协方差的对数行列式，$\log\det\Sigma$）。

我们从两个方面论证这一代理指标的合理性：

1. **几何界：** 从几何角度看，熵衡量不确定性椭球的体积，而积分方差衡量平均平方半径。根据算术-几何平均不等式（以及特征值的性质），约束体积（D-最优性）会隐式地约束迹（A-最优性），因此构成了一个紧致的信息论代理。2. **次模性：** 与方差缩减不同，互信息是一个次模集函数。这一性质至关重要，因为它保证 R-Design 所采用的贪心采集策略能够达到全局最优设计的常数因子近似。

因此，R-EPIG 提供了一种稳健的替代指标，兼具信息论上的严谨性与算法效率。

D.4 引理的证明

我们对 R-Design 估计量的极小极大最优误差上界进行严格推导。证明通过将总CATE风险分解为各分量的估计误差，运用三角不等式将观测基线误差与残差学习误差分离，最后在最优光滑性匹配条件下调用后验收缩速率。

1. 符号与定义 令 $\psi(\hat{\tau}) \coloneqq \|\hat{\tau} - \tau\|^2_{L_2(p)}$ 表示相对于目标协变量密度 $p(\boldsymbol{x})$ 的 PEHE 风险。为简洁起见，我们将正文中引入的有效学习率指数函数定义为：

$$\gamma(\alpha, d) \coloneqq \frac{2\alpha}{2\alpha + d}. \tag{69}$$

我们将处理组 $t \in \{0,1\}$ 的真实潜在结果定义为 $\mu_e(\boldsymbol{x},t)$。真实 CATE 为$\tau(\boldsymbol{x}) = \mu_e(\boldsymbol{x},1) - \mu_e(\boldsymbol{x},0)$。R-Design 估计量定义为 $\hat{\tau}(\boldsymbol{x}) = \hat{\mu}_e(\boldsymbol{x},1) - \hat{\mu}_e(\boldsymbol{x},0)$，其中每个处理组的估计量由第一阶段基准与第二阶段残差之和构成：

$$\hat{\mu}_e(\boldsymbol{x},t) \coloneqq \hat{\mu}_o(\boldsymbol{x},t) + \hat{\delta}(\boldsymbol{x},t). \tag{70}$$

$\hat{}$ 在此，$\hat{\mu}_o$ 由 n_O 观察性样本学习得到，而 δ 由 n_E 实验样本学习得到。

2. 真值的函数分解为严格分析两阶段误差，我们从概念上将真实潜在结果 $\mu_e(\boldsymbol{x},t)$ 分解为一个观察性极限和一个因果残差。令 $\mu_o(\boldsymbol{x},t)$ 表示观察性结果的真实条件期望（即第一阶段的无限数据极限）。该函数反映了相关性和潜在的混杂偏倚。随后，我们将真实结构残差 $\delta(\boldsymbol{x},t)$ 定义为两者之差：

$$\delta(\boldsymbol{x},t) \coloneqq \mu_e(\boldsymbol{x},t) - \mu_o(\boldsymbol{x},t). \tag{71}$$

因此，真值可分解为 $\mu_e(\boldsymbol{x},t) = \mu_o(\boldsymbol{x},t) + \delta(\boldsymbol{x},t)$。关于函数空间的假设：

- 观测基线 $\mu_o(\cdot,t)$ 属于一个霍尔德球 $H^{\alpha_{\mu_o,t}}$，其有效维数为$d_{\mu_o,t}$。
- 残差 $\delta(\cdot,\ t)$ 属于霍尔德球 $H^{\alpha_{\delta,\ t}}$ 其有效维数为 $d_{\delta,\ t}$。

3. 通过特定处理组风险界定条件平均处理效应风险 我们展开 PEHE 风险，并两次应用不等式 $(a-b)^2 \le 2a^2+2b^2$，以解耦各处理臂：

$$\begin{aligned}\psi(\hat{\tau}) &= \|\hat{\tau}-\tau\|_{L_2}^2 \\ &= \|(\hat{\mu}_e(\cdot,1)-\hat{\mu}_e(\cdot,0))-(\mu_e(\cdot,1)-\mu_e(\cdot,0))\|_{L_2}^2 \\ &= \|(\hat{\mu}_e(\cdot,1)-\mu_e(\cdot,1))-(\hat{\mu}_e(\cdot,0)-\mu_e(\cdot,0))\|_{L_2}^2 \\ &\le 2\|\hat{\mu}_e(\cdot,1)-\mu_e(\cdot,1)\|_{L_2}^2+2\|\hat{\mu}_e(\cdot,0)-\mu_e(\cdot,0)\|_{L_2}^2. \end{aligned} \tag{72}$$

只需分析单个处理臂 t 的复合估计量 $\hat{}\mu_e(\cdot,\ t)$ 的收敛性。

4. 解耦第一阶段和第二阶段误差 将估计量的定义 ($\hat{}\mu_e = \hat{}\mu_o + \delta$) 和真实值 ($\mu_e = \mu_o + \delta$) 代入处理臂 t 的误差项。

$$\begin{aligned}\|\hat{\mu}_e(\cdot,t)-\mu_e(\cdot,t)\|_{L_2}^2 &= \|(\hat{\mu}_o(\cdot,t)+\hat{\delta}(\cdot,t))-(\mu_o(\cdot,t)+\delta(\cdot,t))\|_{L_2}^2 \\ &= \|(\hat{\mu}_o(\cdot,t)-\mu_o(\cdot,t))+(\hat{\delta}(\cdot,t)-\delta(\cdot,t))\|_{L_2}^2. \end{aligned} \tag{73}$$

应用不等式 $\|a+b\|^2 \le 2\|a\|^2+2\|b\|^2$，我们将误差分离出来：

$$\|\hat{\mu}_e(\cdot,t)-\mu_e(\cdot,t)\|_{L_2}^2 \le 2\underbrace{\|\hat{\mu}_o(\cdot,t)-\mu_o(\cdot,t)\|_{L_2}^2}_{\text{Term A: Stage 1 Error}}+2\underbrace{\|\hat{\delta}(\cdot,t)-\delta(\cdot,t)\|_{L_2}^2}_{\text{Term B: Residual deviation}}. \tag{74}$$

5. 项 A 的分析（观测基线收敛） 项 A 表示观测学习器的估计误差$\hat{}\mu_o$，该学习器基于 n_O 个样本进行训练，其目标为 μ_o。我们采用Alaa 和 Schaar （2018年，定理2）给出的高斯过程后验收缩率。由于我们假设第一阶段先验光滑度 $\beta_{\mu_o,\ t}$ 与真实光滑度 $\alpha_{\mu_o,\ t}$ （即 $\beta_{\mu_o,\ t} \asymp \alpha_{\mu_o,\ t}$）最优匹配，收缩率可简化为极小极大最优速率。具体而言，对于属于 $H^{\alpha_{\mu_o,\ t}}$ 且有效维数为 $d_{\mu_o,\ t}$ 的函数：

$$\|\hat{\mu}_o(\cdot,t)-\mu_o(\cdot,t)\|_{L_2}^2 \lesssim n_O^{-\gamma(\alpha_{\mu_o,t},d_{\mu_o,t})}. \tag{75}$$

6. 项 B 的分析（残差中的误差传播） 这是关键步骤。项 B 衡量残差学习器 δ 相对于真实结构残差 δ 的偏差 $\tilde{}$。然而，该学习器 $\hat{}\delta$ 是基于经验残差 $y-\hat{}\mu_o(\boldsymbol{x},\ t)$ 进行训练的，实际上是在逼近偏移后的目标$\tilde{}\delta(\boldsymbol{x},\ t)$ $\coloneqq \mu_e(\boldsymbol{x},\ t)-\hat{}\mu_o(\boldsymbol{x},\ t)$。我们通过引入这个中间目标 δ 来分解项 B：

$$\begin{aligned}\|\hat{\delta}(\cdot,t)-\delta(\cdot,t)\|_{L_2}^2 &= \|(\hat{\delta}(\cdot,t)-\tilde{\delta}(\cdot,t))+(\tilde{\delta}(\cdot,t)-\delta(\cdot,t))\|_{L_2}^2 \\ &\le 2\underbrace{\|\hat{\delta}(\cdot,t)-\tilde{\delta}(\cdot,t)\|_{L_2}^2}_{\text{B1: Stage 2 Estimation}}+2\underbrace{\|\tilde{\delta}(\cdot,t)-\delta(\cdot,t)\|_{L_2}^2}_{\text{B2: Bias Propagation}}. \end{aligned} \tag{76}$$

$\tilde{}$*B2*（偏差传播）分析：代入.....的定义 δ 和 δ：

$$\|\tilde{\delta}-\delta\|_{L_2}^2 = \|(\mu_e-\hat{\mu}_o)-(\mu_e-\mu_o)\|_{L_2}^2 = \|\mu_o-\hat{\mu}_o\|_{L_2}^2. \tag{77}$$

该项正是第一阶段误差，其上界由公式75给出。

$\hat{}$*B1*（第二阶段估计）分析：该项表示使用 n_E 个样本时，δ 收敛到其经验目标$\tilde{}\delta$ 的情况。假设第二阶段先验与残差的光滑度相匹配$\alpha_{\delta,\ t}$，则适用标准的高斯过程收缩率：

$$\|\hat{\delta}(\cdot,t)-\tilde{\delta}(\cdot,t)\|_{L_2}^2 \lesssim n_E^{-\gamma(\alpha_{\delta,t},d_{\delta,t})}. \tag{78}$$

$\tilde{}$（注：虽然目标 δ 包含来自 $\hat{}\mu_o$ 的“粗糙”成分，但针对带噪目标的标准非参数回归结果使我们能够将估计误差界定为“信号”类 δ 的速率加上偏差噪声的幅度，而我们已将其分离为 B2）。

7. 最终结果。 将 B1 和 B2 的界代回项 B 的表达式中，然后与项 A 结合到公式74:

$$\begin{aligned}\text{Risk}_t &\le 2(\text{Term A}) + 2(2(\text{B1}) + 2(\text{B2})) \\ &= 2\|\hat{\mu}_o - \mu_o\|^2 + 4\|\hat{\delta} - \tilde{\delta}\|^2 + 4\|\mu_o - \hat{\mu}_o\|^2 \\ &= 6\underbrace{\|\hat{\mu}_o - \mu_o\|^2}_{\lesssim n_O^{-\gamma_{\mu_o}}} + 4\underbrace{\|\hat{\delta} - \tilde{\delta}\|^2}_{\lesssim n_E^{-\gamma_\delta}}.\end{aligned} \tag{79}$$

对处理臂 $t \in \{0, 1\}$ 求和，并将常数（C_1, C_2）及不可约近似误差吸收进去:

$$\psi(\hat{\tau}) \le C_1 \max_t n_E^{-\gamma(\alpha_{\delta,t}, d_{\delta,t})} + C_2 \max_t n_O^{-\gamma(\alpha_{\mu_o,t}, d_{\mu_o,t})} + \epsilon_{\text{approx}}. \tag{80}$$

这证实，总误差由相应收敛速率之和主导，从而验证了残差更平滑（$\alpha_\delta > \alpha_{\mu_o}$）且 n_O 较大时的效率增益。 □

D.5 命题 2 的证明

在本部分中，我们对第 5.3 节所述的信息冗余分解进行形式化推导。我们证明，最大化基于参数的互信息，等价于最大化目标估计量信息与一个非负干扰项之和。

证明。 该证明依赖于应用于实验结果 y、模型参数 θ 以及残差对比预测 $\tau_\delta(\boldsymbol{x}^*)$ 联合分布的互信息链式法则，其中目标点为 $\boldsymbol{x}^*$。我们按两种等价顺序展开联合互信息 $\mathrm{I}(y; \theta, \tau_\delta(\boldsymbol{x}^*))$:

$$\mathrm{I}(y; \theta, \tau_\delta(\boldsymbol{x}^*)) = \mathrm{I}(y; \theta) + \mathrm{I}(y; \tau_\delta(\boldsymbol{x}^*) \mid \theta) \tag{81}$$

$$= \mathrm{I}(y; \tau_\delta(\boldsymbol{x}^*)) + \mathrm{I}(y; \theta \mid \tau_\delta(\boldsymbol{x}^*)). \tag{82}$$

步长 1：确定性依赖。关键在于，预测 $\tau_\delta(\boldsymbol{x}^*)$ 是参数 θ（即 $\tau_\delta(\boldsymbol{x}^*) = f(\boldsymbol{x}^*; \theta)$）的确定性函数。因此，给定 θ后，$\tau_\delta(\boldsymbol{x}^*)$中不再存在不确定性。这意味着式81 中的条件互信息为零:

$$\mathrm{I}(y; \tau_\delta(\boldsymbol{x}^*) \mid \theta) = 0. \tag{83}$$

步长 2：等式转换与整理。将这个零项代回式 81 并令其等于式 82，可得:

$$\mathrm{I}(y; \theta) = \mathrm{I}(y; \tau_\delta(\boldsymbol{x}^*)) + \mathrm{I}(y; \theta \mid \tau_\delta(\boldsymbol{x}^*)). \tag{84}$$

步长 3：对目标总体取期望。标准的 R-EPIG 目标函数定义为对目标总体 D_O 取期望。将所有内容以当前数据集 D 为条件，并取期望 $\mathbb{E}_{\boldsymbol{x}^* \sim p_{\text{tar}}(\boldsymbol{x})}$，注意到参数信息项 $\mathrm{I}(y; \theta \mid \mathrm{D})$ 与具体的评估点 $\boldsymbol{x}^*$ 无关。因此:

$$\mathrm{I}(y; \theta \mid \mathcal{D}) = \mathbb{E}_{\boldsymbol{x}^*}[\mathrm{I}(y; \tau_\delta(\boldsymbol{x}^*) \mid \mathcal{D})] + \mathbb{E}_{\boldsymbol{x}^*}[\mathrm{I}(y; \theta \mid \tau_\delta(\boldsymbol{x}^*), \mathcal{D})]. \tag{85}$$

对应的各项即为命题2中的分解:

$$\text{Residual-BALD} = \text{R-EPIG} + \text{Expected Redundancy}. \tag{86}$$

由于互信息始终非负，因此冗余项为 ≥ 0。证明至此结束。

□

E 更多实验结果

在本部分中，我们提供一组全面的补充实验结果，以补充主要发现。首先，我们展示了合成和半合成（IHDP、ACTG-175）基准上主要实验的更多性能曲线。其次，我们开展了一系列消融研究，以分析 R-Design 框架（包括 R-EPIG 和 TSR策略）的鲁棒性。这些研究评估了初始随机起点、采集批次大小以及未标注池规模变化所带来的影响。

E.1 合成数据集与消融实验

E.1.1 模拟的更多结果

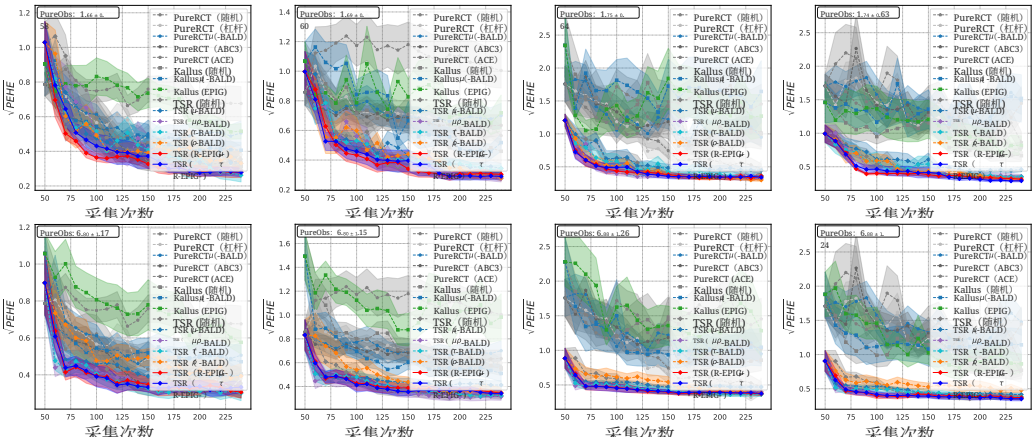


图9： √PEHE 单变量模拟设置比较。第一行展示编号为 1–4 的模拟设置结果，第二行展示编号为 5–8 的结果。

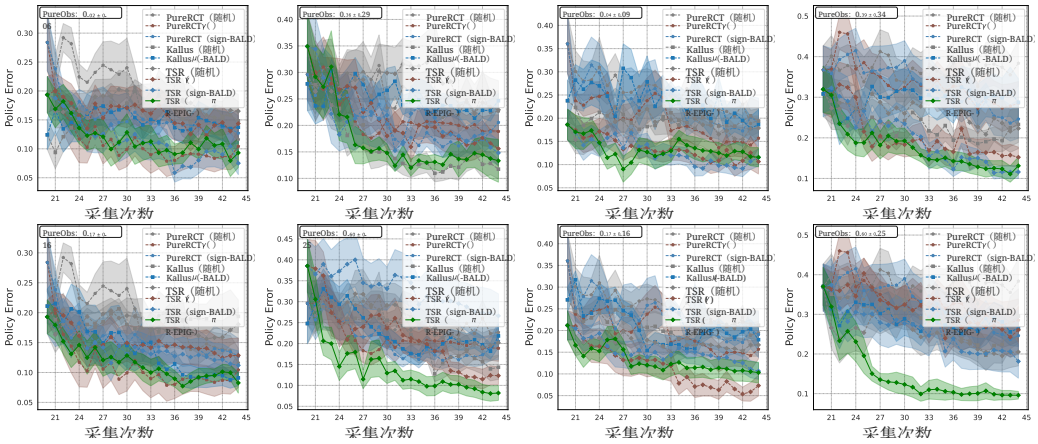


图10：**平均策略误差（APE）收敛**。对八个单变量模拟基准进行比较。顶行展示情景 1–4 的结果，底行展示情景 5–8 的结果。APE 越低，识别最优治疗臂的准确度越高。

单变量情形

效应估计精度（CATE）。我们首先评估 R-Design 在第一项任务上的表现，即处理效应估计的准确性。图9 报告了 √来自Dimitriou 等人（2024年）的八个单变量基准上的 PEHE 收敛情况。

结构优势。 结果充分验证了所提出 TSR 架构的效率。在所有场景中，基于 TSR 的方法始终优于 PureRCT（Cha&Lee，2025年）和 Kallus（Kallus 等人，2018年）基线。表现最佳的 TSR 配置相较于 PureRCT-Random 可将估计误差降低 9–63%，相较于 Kallus-Random 可降低 10–66%。在复杂混杂条件下，这一性能差距尤为明显（例如模拟 4，相较于 PureRCT 提升了 62.7%）。此外，TSR 方法还表现出显著更高的稳定性：各次运行的平均标准差为 0.075，而 PureRCT 和 Kallus 分别为 0.231 和 0.302。这表明，以观测先验为基础开展试验学习，能够有效缓解 RCT样本有限所引起的方差。

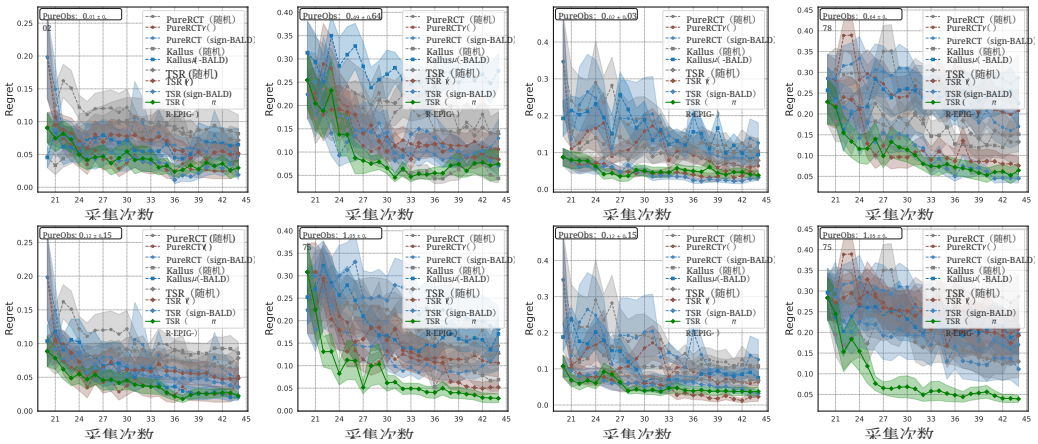


图11：**策略遗憾收敛**。在八个单变量模拟基准上的比较。上行展示场景1–4的结果，下行展示场景5–8的结果。R-EPIG-π在大多数任务中最快收敛至零遗憾。

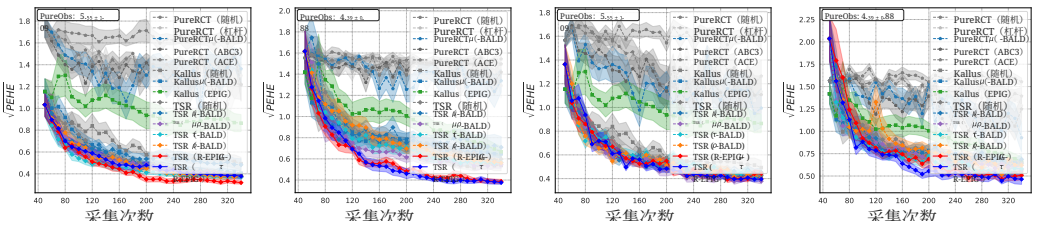


图12：不同轨迹模型在 dim 6 数据集上的 PEHE 比较。从左到右：正态数据集上的 CMGP、严重协变量偏移下的 CMGP、正态数据集上的 NSGP、严重协变量偏移下的 NSGP。

采集效率。 在 TSR框架中，所提出的 R-EPIG-τ 采集策略展现出强大的样本效率，在 8 次模拟中的 5 次排名第一。在模拟1中，仅经过 190 次采集，R-EPIG-τ就将初始估计误差降低了 70% 以上，而 PureRCT 的降幅为 55%。这证实，主动针对残差曲面中的认知不确定性，相较于随机采样能够实现更快的收敛。

策略优化的有效性。接下来，我们评估决策表现，其目标是正确识别最优处理策略 $\pi^*(x) = \mathbb{I}[\tau(x) > 0]$。图10 和11分别报告了 APE 和平均遗憾的收敛情况。

决策感知学习。结果凸显了使采集目标与下游决策任务相一致的重要性。所提出的决策感知采集函数 R-EPIG-π，在所有方法中实现了最低的平均策略误差（0.105）和遗憾（0.042）。与 PureRCT-Random 和 Kallus-Random 相比，R-EPIG-π 分别将平均遗憾降低了 72% 和 61%。这一优势源于它能够将采样集中在决策边界附近（$\tau(x) \approx 0$）；而以估计为导向的策略可能会耗费预算，在最优决策已明确的区域消除不确定性。

总体表现。 决策表现排名前四的方法全部基于TSR（包括TSR-sign-BALD和TSR-γ），凸显了残差学习为策略优化提供了更优越的基础。总体而言，与PureRCT基线相比，TSR方法的策略误差降低了20–61%，这表明，校正有偏差的观测模型，对于决策而言，远比从头开始学习策略有效。

多变量情形

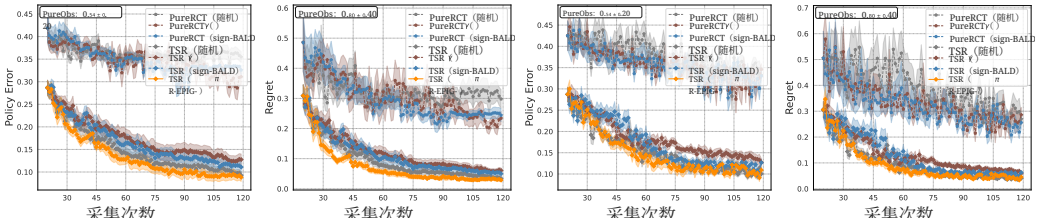


图13：不同轨迹模型下，在 dim 6 数据集上的决策表现比较。从左至右：CMGP（APE）、CMGP（AR）、NSGP（APE）、NSGP（AR）。

多变量CATE估计。 图 12 展示了在6维多变量基准上，采用CMGP和NSGP作为试验结果模型时，标准协变量偏移和严重协变量偏移两种设置下PEHE的收敛情况。在所有配置中，结合R-EPIG采集函数的TSR方法始终达到最低估计误差。在标准数据集上，采用CMGP的R-EPIG-μ 取得最佳最终PEHE，达到0.319；相较于PureRCT-Random（1.220），误差降低了73.8%，较Kallus-Random（0.493）改善了35.3%。采用NSGP作为试验模型时，R-EPIG-τ 表现最佳（0.396），较PureRCT和Kallus分别改善了68.7%和19.8%。

在严重协变量偏移下，随机对照试验与目标总体之间的分布失配更加严重，R-EPIG 的优势也愈发明显。对于 CMGP，R-EPIG-τ 的 $\sqrt{\text{PEHE}}$ 达到 0.376，与 PureRCT（1.497）相比，误差降低了 74.9%；与 Kallus（0.538）相比，误差降低了 30.0%。对于 NSGP，R-EPIG-τ 达到 0.465，分别优于 PureRCT 和 Kallus 66.7% 和 13.5%。值得注意的是，R-EPIG-μ 和 R-EPIG-τ 在协变量偏移下均保持稳健，始终优于所有基线方法。

多变量决策。 图 13 报告了六维基准上的决策性能，该性能通过CMGP和NSGP试验模型的APE和AR进行衡量。与PureRCT基线相比，基于TSR的方法显著降低了策略误差和遗憾。在CMGP模型下，R-EPIG-π 取得了最低的策略误差（0.088）和遗憾（0.029），相较于PureRCT-Random（0.355和0.298），APE和AR分别降低了75.1%和90.4%。在NSGP模型下，TSR-Random取得了最低的策略误差（0.101），而R-EPIG-π 取得了最低的遗憾（0.041），相较于PureRCT分别提升了69.7%和84.9%。

总体而言，这些结果表明：(i) TSR框架通过有效利用观测数据，显著提升了高维CATE估计和治疗决策；(ii) R-EPIG采集函数在不同试验模型下均能带来稳定的增益；(iii) 这种优势在决策导向指标上尤为明显，与纯基于随机对照试验的方法相比，R-EPIG-π 将遗憾降低了85%以上。

> **要点：**在单变量和多变量合成基准上，TSR框架通过利用观测数据修复偏差，持续实现更优的条件平均处理效应估计和策略学习性能。R-EPIG采集函数通过明确针对认知不确定性进一步提高样本效率，而决策感知变体（R-EPIG-π）在下游处理决策为主要目标时带来最大增益。这些结果表明，将偏差校正与主动试验学习分离，为自适应实验提供了有原则且有效的基础。

E.1.2 消融实验

核类型消融。图14 比较了在三种高斯过程核——径向基函数（RBF）、Matérn 和有理二次（RQ）——下，TSR 与 R-EPIG 采集方法在以 CMGP 为试验结果模型的六维数据集上的表现。在两种 R-EPIG 变体中，RBF 核始终表现最佳。使用 R-EPIG-μ时，RBF 的最终$\sqrt{\text{PEHE}}$ 最低，为~~0.319~~，较 Matérn（0.359）低 11.1%，较 RQ（0.385）低 17.1%。归一化曲线下面积（0.454，而 Matérn 为 0.639，RQ 为 0.626）也体现了这一优势，表明

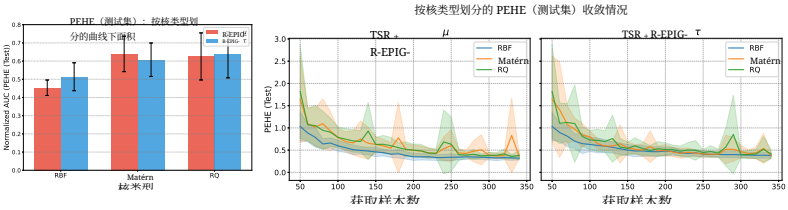


图14：比　　较条件平均处理效应估计性能与　　不同的核。

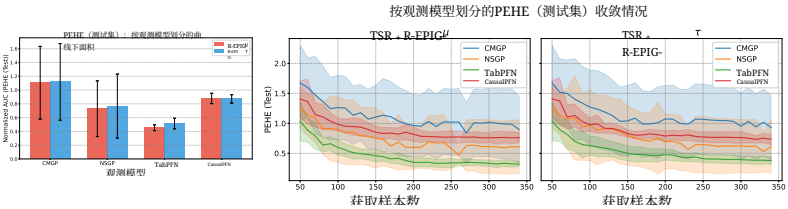


图15：不同观测模型下 CATE 估计性能的比较。

fast采集过程中收敛速度更快。R-EPIG-τ 也呈现类似趋势，其中 RB　　F (0.378）虽优势不大，但始终优于 Matérn（0.381）和 RQ（0.411）。

除了准确性之外，与 RQ（std = 0.095）相比，RBF 在不同随机种子间的方差显著更低（R-EPIG-μ 的 std = 0.040），这表明其优化过程更加稳定。总体而言，尽管 R-EPIG 采集在不同核选择下均表现有效，但对于本文考虑的合成数据生成过程，更简单的径向基函数核在准确性、收敛速度和稳定性之间实现了最佳权衡。

观测模型消融。 图15 使用 R-EPIG 采集，在 TSR 框架中评估不同观测模型（CMGP、NSGP、TabPFN 和 CausalPFN）的影响。所有实验均在六维数据集上进行，并以 CMGP 作为试验结果模型。TabPFN 显著优于所有其他观测模型。使用 R-EPIG-μ时，TabPFN 的最终 $\sqrt{\text{PEHE}}$ 为 0.319，较 CMGP（0.897）、NSGP（0.608）和 CausalPFN（0.762）分别提升 64.4%、47.5% 和 58.1%。对于 R-EPIG-τ，提升幅度同样显著：TabPFN（0.378）较 CMGP、NSGP 和 CausalPFN 分别高出 59.4%、38.6% 和 48.6%。

归一化曲线下面积进一步凸显了TabPFN的优势：在R-EPIG-μ下，TabPFN的曲线下面积为0.454，而CMGP为1.106（降低了2.4× ），表明收敛速度显著更快。此外，与基于高斯过程的观测模型（CMGP：0.559；NSGP：0.448）相比，TabPFN的方差显著更低（标准差 = 0.040），表明偏差估计更为稳健。这些结果表明，TabPFN的上下文学习为两阶段回归提供了更强的观测先验，从而实现更准确的偏差校正和更有效的主动采集。

池大小消融。 图 16 研究候选池大小对采集性能的影响，在以CMGP为试验模型的六维数据集上评估了从500到2500的池大小。R-EPIG采集函数在所有池大小上均持续优于随机采样和基线方法。R-EPIG 变体在全部五个池大小上均取得最佳性能：R-EPIG-μ在池大小为500、1000、1500和2500时胜出，而R-EPIG-τ在池大小为2000时表现最佳。在各种设置下，最佳TSR方法相较PureRCT-Random提升71–76%，相较Kallus-Random提升66–83%。

值得注意的是，随着池大小增加，TSR方法仍保持稳定，R-EPIG-μ 的 $\sqrt{\text{PEHE}}$值介于0.319和0.401之间。相比之下，基线方法的表现会随着候选池扩大而变差：Kallus-Random的性能显著下降（池大小为500时为1.09，增至2500时为1.94），PureRCT-Random的表现也有所恶化（从1.29升至1.41）。这些结果表明，随着搜索空间扩大，智能采集的价值日益凸显，而随机选择难以随着池大小增加而有效扩展。

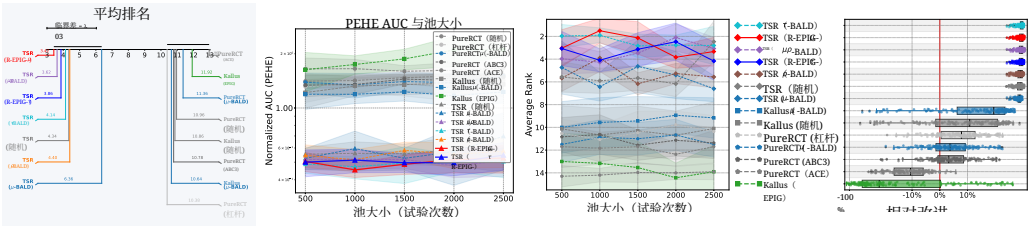


图16：不同池大小下的CATE估计性能比较。

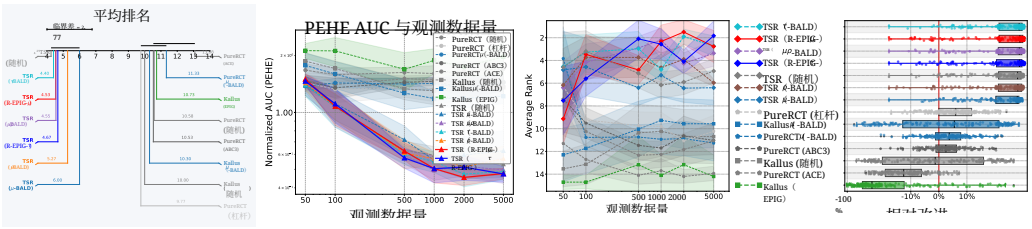


图17：不同观测样本规模下的条件平均处理效应估计性能比较。

观测数据规模消融实验。图17 考察了可用观测数据量对TSR性能的影响，观测样本量从50变化至5000。随着观测数据增加，TSR性能显著提升。R-EPIG-μ 将 $\sqrt{\text{PEHE}}$从50个样本时的1.410降至5000个样本时的0.336，相当于降低了76.2%。这一趋势证实，两阶段框架能够有效利用更大的观测数据集来改进偏差校正。

重要的是，只有当观测模型足够准确时，定向采集才会带来收益。在观测样本规模很小（50–100）时，TSR-Random 的表现优于 R-EPIG，因为观测模型给出的不确定性估计并不可靠。然而，随着数据量增加，R-EPIG 逐渐占据优势，并在 6 种设置中的 4 种（500、1000、2000、5000）取得最佳表现。当样本数为 5000 时，R-EPIG-τ 的 $\sqrt{\text{PEHE}}$最低，为 0.309。总体而言，当观测数据充足（500+样本）时，采用 R-EPIG 的 TSR 相比 PureRCT 可提升 69–76%，相比 Kallus 可提升 57–74%。

批次大小消融。 图 18 分析了采集批次大小的影响，评估范围为 2 到 30。R-EPIG 采集函数在 6 种批次大小中的 5 种下优于随机采样，且在小到中等批次大小时收益最大。批次大小为 10 时，R-EPIG-μ 取得了最佳 $\sqrt{\text{PEHE}}$，为 0.319，相比随机采样（0.379）提升了 15.7%。批次大小为 2、5 和 20 时，提升幅度为 8.7% 至 13.1%。

当批次大小超过 20 时，定向采集的优势会逐渐减弱。批次大小为 30 时，随机采样的表现略优于 R-EPIG（0.371 对比 0.385），这表明选择大批次会削弱信息论准则的有效性。总体而言，R-EPIG 在序贯或小批次场景中表现最佳，因为此时可以充分发挥精细化样本选择的优势。在批次大小为 2–20 的范围内，两个 R-EPIG 变体的表现均保持稳定（0.319–0.385），表明该方法对这一超参数具有鲁棒性。

试验模型消融。 图。19（左侧和中间）比较了使用 TabPFN 作为观测模型时，五种试验结果模型——CMGP、NSGP、BART、BCF 和 CMDE——的 TSR 性能。基于 GP 的试验模型始终优于基于树的模型和集成替代模型。CMGP 的性能最佳，R-EPIG-μ 达到 $\sqrt{0.319}$ 的 PEHE，其次是 NSGP（使用 R-EPIG-τ时为 0.396）。相比之下，BART（0.681）、BCF（0.799）和 CMDE（0.812）的估计误差显著更高。这一差距凸显了校准良好的不确定性估计的重要性；基于 GP 的模型能够提供这种估计，而它对于有效的信息论采集至关重要。

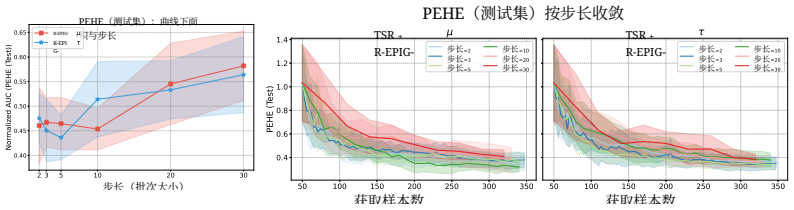


图18：不同批次大小下的CATE估计性能比较。

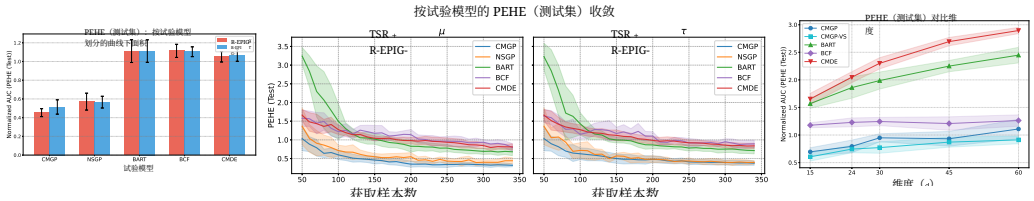


图19：不同试验模型下的CATE估计性能比较。

总体而言，对于基于 GP 的试验模型，R-EPIG 优于随机采样（CMGP 提升 15.7%，NSGP 提升 7.1%）；而对于基于树的模型，由于其不确定性估计的可靠性较低，提升幅度有限或不存在。

高维可扩展性。 图 19 （右）使用CMGP作为试验模型，评估了维度从15增加到60时TSR的可扩展性。R-EPIG-μ 在全部五个维度上均取得了最佳性能，表明该信息论采集方法在更高维度的设置下仍然有效。性能平稳下降： $\sqrt{\text{PEHE}}$从15维度时的0.462增加到60维度时的0.691，尽管维度增加了，仍仅中等幅度上升49%； 4× 维度增加。这些结果证实，采用R-EPIG采集的TSR能够稳健地扩展到中等高维问题。

> **要点：** 消融研究表明，R-EPIG的有效性在不同核函数选择、候选池大小、批次大小以及维度中等幅度增加的情况下均具有稳健性，同时最受益于校准良好的不确定性和信息足够充分的观测模型。只有在不确定性估计不可靠时（例如观测数据集极小或试验模型校准不佳），性能提升才会减弱，这凸显了不确定性质量在信息论采集中的核心作用。总体而言，这些发现证实，在现实的建模和系统设计选择下，R-EPIG能够提供一致且可预测的改进。

E.2 半合成数据集

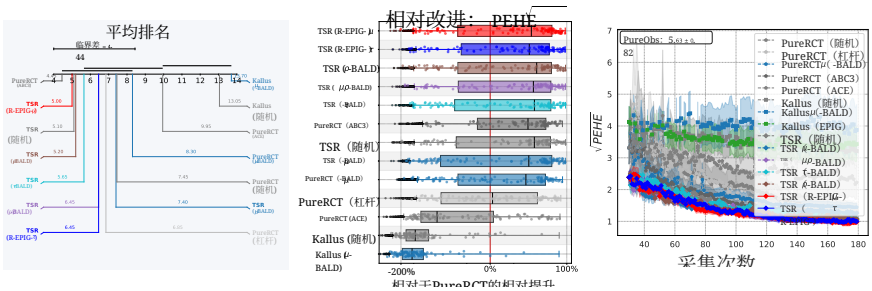


图20：IHDP 数据集上的CATE估计性能比较。

IHDP 条件平均处理效应估计。图20报告了 IHDP 数据集上的条件平均处理效应估计结果。该数据集是基于一项真实世界的婴儿健康研究构建、广泛使用的半合成基准。在 CMGP 作为

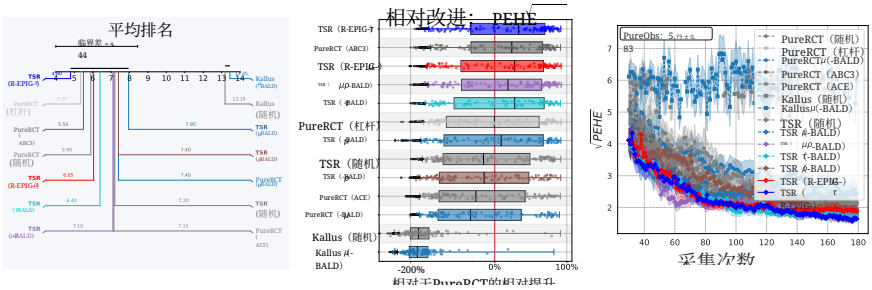


图 21：ACTG-175 数据集上的 CATE估计性能比较。

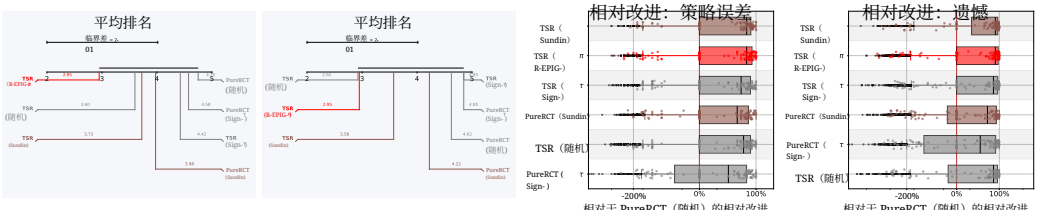


图 22：IHDP 数据集上的策略估计性能比较。

试验结果模型的情况下，TSR 与 R-EPIG-μ 结合取得了最佳表现，PEHE 为$\sqrt{0.980}$，紧随其后的是 R-EPIG-τ（0.999）。相较于 PureRCT-Random（1.297），最佳 TSR 配置将估计误差降低了24.5%。相较于 Kallus-Random，提升幅度显著更大，达到69.5%（3.213），凸显了通过残差学习明确校正观察性偏差的优势。

当使用 NSGP 作为试验模型时，TSR-Random 的表现具有竞争力（1.032），这表明定向采集带来的额外收益取决于试验模型不确定性的校准质量。尽管如此，在两种试验模型下，基于 TSR 的方法始终优于 PureRCT 和 Kallus 基线，证明了在真实的非合成环境中利用观测数据进行偏差校正的有效性。

IHDP决策。 图 22 使用 APE 和 AR 评估 IHDP 上的治疗策略学习。采用 CMGP 时，所提出的决策感知采集函数 R-EPIG-π 取得了最低的策略误差（0.018）和遗憾（0.015）；与 PureRCT-Random 相比，策略误差（0.034）降低了 47.0%，遗憾（0.040）降低了 61.9%。采用 NSGP 时，TSR-Random 和 R-EPIG-π 的策略误差相当（0.020），而 TSR-Random 的遗憾略低（0.012 对 0.017）。

总体而言，这些结果验证了以决策为导向的采集在真实世界数据上仍然有效，相较于纯粹基于随机对照试验的方法取得了显著提升，同时也表明，收益幅度取决于试验模型所提供的不确定性估计的可靠性。

ACTG-175条件平均处理效应估计。 图 21 报告了 ACTG-175 数据集上的条件平均处理效应估计性能。该数据集源自一项艾滋病临床试验，其处理效应更加复杂且具有异质性。使用 CMGP 时，TSR 与 R-EPIG-τ 结合取得了最佳 $\sqrt{\text{PEHE}}$（1.647），比 PureRCT-Random（1.908）提升了 13.7%。使用 NSGP 时，R-EPIG-τ（1.778）略优于 TSR-Random（1.880），表明定向采集仍有益处，不过其优势幅度小于在 IHDP 上的表现。

值得注意的是，基于Kallus的方法在该数据集上的表现不佳：Kallus-μ 的 $\sqrt{\text{PEHE}}$达到 5.993，误差是基于TSR的方法的三倍以上。这一差距凸显了仅依赖试验数据、却不明确校正观察性偏差的方法存在局限，尤其是在具有挑战性的现实场景中。

ACTG-175 决策。 图 23 展示了 ACTG-175 上的治疗策略学习结果。使用 CMGP 时，决策导向TSR方法实现了最强的总体表现。TSR-sign-BALD

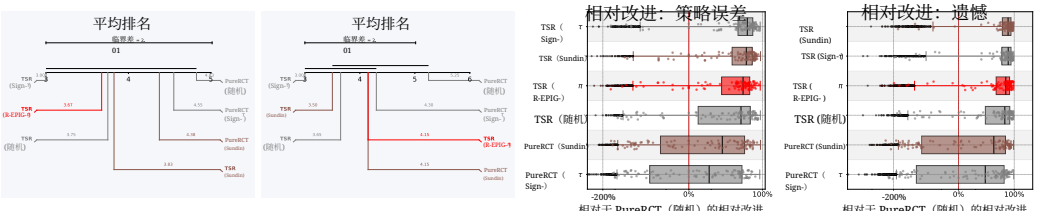


图 23：ACTG-175 数据集上的策略估计性能比较。

取得了最低的策略误差（0.138）和遗憾（0.113），紧随其后的是 TSR-γ（PE：0.139，Reg：0.122）和 R-EPIG-π（PE：0.138，Reg：0.130）。所有 TSR变体均显著优于 PureRCT-Random（PE：0.190，Reg：0.195），策略误差提高了 27–42%，遗憾提高了 33–42%。使用 NSGP 时，TSR-sign-BALD 再次取得最佳总体表现（PE：0.131，Reg：0.116）。

综上，这些结果表明，决策感知采集函数（sign-BALD、γ和 R-EPIG-π）能够持续提升真实世界数据集上的治疗策略学习效果。在 IHDP 和 ACTG-175 数据集上，基于TSR的方法均稳定优于PureRCT基线，这证实了残差学习与有原则的采集相结合，能够在实践中实现更有效、更稳健的决策。

> **要点：** 在真实世界基准（IHDP 和 ACTG-175）上，即使没有预言机假设，基于TSR的方法在条件平均处理效应估计和治疗决策方面也始终优于纯随机对照试验基线。当不确定性经过良好校准时，决策感知采集函数带来的提升最为显著；而在噪声更大的情形下，TSR-Random 仍是稳健的备选方案。这些结果表明，残差学习与有原则的采集相结合，不仅在受控的合成环境中有效，也适用于真实世界的因果决策问题，具有实用性和可靠性。

F 讨论：JOINTMODEL 还是 TSR?

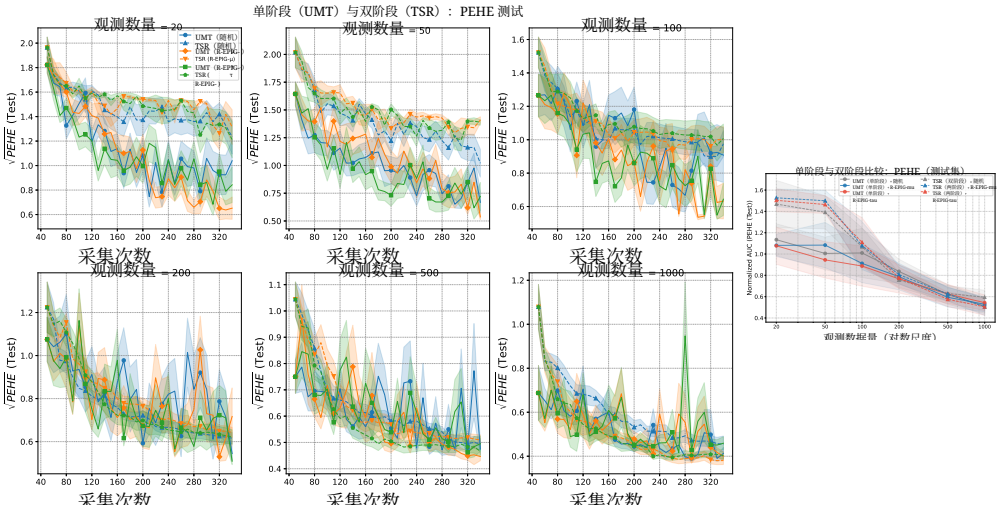


图 24：两种不同策略下的 CATE估计性能比较。

在正文中，我们使用 TSR 框架介绍了 R-Design。不过，还有一种替代策略：在统一的概率框架中联合建模观测数据和实验数据，这与 CausalICM 的思路非常相似（Dimitriou 等人，2024年）。然而，CausalICM 仅对各组内部的相关性建模，即分别使用两个 CMGP 对照组和处理组建模。这是为了利用大量观测数据中的信息来帮助预测

随机对照试验数据。在这一部分，我们讨论一种更一般的情形：使用一个模型对所有关注的输出进行联合建模，并正式定义这种统一多任务（UMT）架构，同时从计算复杂度、误差传播和样本效率等方面与 TSR 进行全面比较。

F.0.1 UMT 框架

与 R-Design 的序贯方法不同，R-Design 会在第 2 阶段的自适应实验设计开始前冻结观测模型参数，而 UMT 方法则将观测结果和实验结果视为单个联合训练的高斯过程的输出。

表述。 我们在扩展输入空间 $\mathrm{X} = \mathrm{X} \times \{0, 1\} \times \{s_o, s_e\}$ 上构建一个多任务高斯过程，其中 s 表示数据源。联合模型 $f_{\text{joint}}(\boldsymbol{x}, t, s)$ 捕捉有偏观测源与无偏实验源之间的相关性。核函数通常采用内禀共区域化模型（ICM）定义：

$$k_{\text{joint}}((\boldsymbol{x}, t, s), (\boldsymbol{x}', t', s')) = k_{\text{input}}((\boldsymbol{x}, t), (\boldsymbol{x}', t')) \otimes \mathbf{B}_{s,s'}, \tag{87}$$

其中 $\mathbf{B}$ 是一个共区域化矩阵，用于捕捉来源 s_o 与 s_e之间的协方差。在此框架中，主动采集步骤包括：1. 在 $\mathrm{D}_O \cup \mathrm{D}_E^{(k)}$ 上训练全局高斯过程。2. 使用实验任务 $f(\cdot, \cdot, s_e)$ 的后验分布计算采集分数（例如 R-EPIG）。3. 观测到新样本 y_{new} 后更新整个 模型（包括来源间的相关性）。

F.0.2 理论比较

在单阶段（UMT）与双阶段（TSR）之间进行选择，涉及信息流与计算可行性之间的根本权衡。

1. 信息流与误差传播。

• **UMT（单阶段）：** 支持双向信息流。实验数据可以追溯性地更新对观测参数的认知。在**低数据量情形**下，这一做法具有优势，因为观察性数据集 D_O 太小，无法形成可靠的先验。在这种情况下，联合训练可以避免“冻结误差”问题：第一阶段模型表现不佳，会永久性地限制残差学习器。

• **TSR（双阶段）：** 施加了严格的结构约束。通过固定 $\hat{}\mu_o$，它迫使阶段 2 模型将所有差异都解释为残差。这样可以隔离实验信号，避免当 n_O 较大时，实验信号被观测似然淹没。

2. 计算复杂度。这是主动学习的关键瓶颈。

• **UMT：** 计算成本随总数据量呈立方增长：$\mathrm{O}((n_O + n_E)^3)$。对于大规模观察性研究（例如 $n_O > 10^4$），在每一步重新训练联合模型并计算采集分数，在计算上代价过高。

• **TSR：** 主动循环仅随实验数据量增长：$\mathrm{O}(n_E^3)$。海量观测数据只需离线处理一次。这使得 TSR 可扩展至大规模现实场景。

F.0.3 实证分析

我们通过一项比较 UMT 与 TSR 的消融研究来验证这些理论洞见。我们在固定实验预算的同时改变观测数据量 n_O，并报告 PEHE性能。

结果讨论。 表5 中的结果与我们的理论分析完全一致：

1. **低数据量场景（$n_O < 200$）：** UMT 方法的表现显著优于 TSR。当观测数据稀缺时，TSR 中的第一阶段估计器非常不准确。由于 TSR 会冻结这个“垃圾”模型，第二阶段学习器面临着学习高度复杂且高方差的残差曲面这一艰巨任务，以进行补偿。UMT 通过联合平滑这两个来源避免了这一问题，并利用实验数据稳定观察数据估计值。

表 5：不同观测数据量（n_O）下，单阶段（UMT）和双阶段（TSR）方法的标准化 PEHE（测试集）曲线下面积比较。两种方法均采用 R-EPIG-μ 采集函数。

方法	**观测数据量（n_O）**					
	$n_O = 20$	$n_O = 50$	$n_O = 100$	$n_O = 200$	$n_O = 500$	$n_O = 1000$
UMT（单阶段）	**1.08**	**1.08**	**0.91**	**0.79**	**0.59**	0.53
TSR（双阶段）	1.53	1.50	1.08	0.81	0.63	**0.50**

2. **过渡与交叉：** 随着 n_O 的增加，TSR 中的第一阶段模型变得可靠，性能差距迅速缩小。“残差”变得更加简单、平滑（符合引理1中的结构假设）。

3. **高数据量区间（$n_O \geq 1000$）：** 在 $n = 1000$ 处出现交叉，此时 TSR（0.50）超过 UMT（0.53）。这凸显了两阶段架构的“偏差隔离”优势。在联合模型中，随着 n_O 占据主导，似然函数可能会优先拟合海量观测数据，而非稀疏的实验点，从而可能削弱偏差校正。TSR 将观测模型视为固定锚点，确保实验预算严格用于细化差异，从而在具备强先验时获得更优的渐近性能。

结论。 尽管 UMT 在数据稀缺场景中较为稳健，但就本文的问题设定而言，R-Design（TSR）是最优选择；该设定假设我们可以获取大规模、预先存在的观测数据集（$n_O \gg n_E$）。TSR 通过有效地将偏差估计与残差校正解耦，提供了更优的可扩展性与渐近精度。