MULTI-OBJECTIVE MULTI-FIDELITY BAYESIAN OPTIMIZATION WITH CAUSAL PRIORS

Anonymous authors
Paper under double-blind review

ABSTRACT

Multi-fidelity Bayesian optimization (MFBO) accelerates the search for the global optimum of black-box functions by integrating inexpensive, low-fidelity approximations. The central task of an MFBO policy is to balance the cost-efficiency of low-fidelity proxies against their reduced accuracy to ensure effective progression toward the high-fidelity optimum. Existing MFBO methods primarily capture associational dependencies between inputs, fidelities, and objectives instead of causal mechanisms, and can perform poorly when lower-fidelity proxies are poorly aligned with the target fidelity. We propose RESCUE (REducing Sampling cost with Causal Understanding and Estimation), a multi-objective MFBO method that incorporates causal calculus to systematically address this challenge. RESCUE learns a structural causal model capturing causal relationships between inputs, fidelities, objectives, and constraints, and uses it to construct a probabilistic multi-fidelity (MF) surrogate that encodes intervention effects. Exploiting the causal structure, we introduce a value-of-causal-information hypervolume knowledge-gradient acquisition strategy to select input–fidelity pairs that balance expected feasible multi-objective improvement and cost. We show that RESCUE improves sample efficiency over state-of-the-art MF optimization methods on synthetic and real-world problems in robotics, and machine learning (AutoML).

1 INTRODUCTION

Optimizing expensive black-box functions without analytical forms or derivative information is a ubiquitous problem. Multi-Objective Bayesian Optimization (MOBO) addresses this problem sample-efficiently (Hernández-Lobato et al., 2016) by iteratively maximizing an acquisition function derived from a probabilistic surrogate model (typically a Gaussian Process (GP)) to select candidate inputs for intervention. The selected inputs are evaluated on the true objective functions, and the surrogate is updated to improve objective estimation. However, executing interventions on the true objective functions is costly, as each intervention can involve extensive computation or expensive real-world experiments. In practice, multiple Information Sources (IS) with varying numerical resolutions are often available to approximate the objective functions. Interventions at low fidelities typically incur lower cost but provide less accurate approximations than interventions at higher fidelities. For example, tuning parameters of a mobile robot can require longer evaluation time subject to safety and operational constraints but can be accelerated using lower-fidelity simulations such as *Gazebo* (Hossen et al., 2025). Practical deployments must also satisfy black-box constraints, i.e. safety limits in robot navigation or resource budgets in AutoML, while infeasible evaluations consume the same scarce budget. A key challenge is optimizing the trade-off between the cost and accuracy of MF interventions to enable informed decision-making. This challenge intensifies when auxiliary IS poorly approximate the target fidelity. In such cases, single-fidelity BO (SFBO), which uses only the primary IS, can outperform MFBO methods under a fixed budget (Mikkola et al., 2023).

In this paper, we present RESCUE, a multi-objective multi-fidelity optimization method integrating causal inference into MOBO. RESCUE builds upon the literature on constrained causal BO (Aglietti et al., 2023), causal knowledge transfer for BO (Hossen et al., 2025), and MF optimization (Mikkola et al., 2023; Poloczek et al., 2017). Specifically, RESCUE constructs a causal model over configuration options, fidelities, objectives, constraints, and intermediate variables from observational data (e.g., system logs) or a domain-provided directed acyclic graph (DAG), a weaker requirement than causal BO methods that assume a known graph (Aglietti et al., 2020; 2023). This causal model provides the

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

causal prior for an MF multi-output GP surrogate, henceforth the MF Causal GP (MF-CGP). The MF-CGP models objectives and constraints across fidelities. RESCUE selects configuration–fidelity pairs by maximizing the Value-of-Causal-Information Hypervolume Knowledge Gradient, an extension of HVKG (Daulton et al., 2023) that learns how much to trust the causal prior. The optimizer exploits cross-fidelity transfer while regularizing it with causal information beyond observed correlations. Our contributions are as follows:

- We develop a constrained multi-objective multi-fidelity optimization method that integrates causal information into the MF-MOBO framework
- We propose the multi-fidelity Value-of-Causal-Information Hypervolume Knowledge Gradient (VoCI-HVKG), a cost-aware acquisition strategy that learns how much to trust the causal prior when selecting an optimal input–fidelity pair while accounting for feasibility.
- Theoretically, we establish a performance bound independent of the cross-fidelity bias magnitude for a bounded causal prior error.
- We demonstrate substantial gains with RESCUE over state-of-the-art MF-MOBO and MOBO methods on synthetic and real-world MF problems in robotics, and AutoML.

2 Related Works

MFBO reduces optimization cost using hierarchical methods when IS can be ranked by fidelity (Kandasamy et al., 2016; Li et al., 2020; Takeno et al., 2020; Moss et al., 2021; Tighineanu et al., 2022) and via non-hierarchical methods otherwise (Lam et al., 2015; Poloczek et al., 2017; Wu et al., 2023). Although hierarchical methods may propagate errors from lower-fidelities (Wu et al., 2023), we do not claim that non-hierarchical methods are inherently superior or inferior; rather, the choice is application specific (Zhou et al., 2023). Our MF-CGP jointly learns cross-fidelity correlations, while the causal model lets experts encode domain knowledge. MFBO methods address both multi-objective (Irshad et al., 2024; Daulton et al., 2023; Belakaria et al., 2020; He et al., 2022) and single-objective (Mikkola et al., 2023; Li et al., 2020; Takeno et al., 2020; Poloczek et al., 2017) optimization. We refer the interested reader to Daulton et al. (2023, Section 3) and Belakaria et al. (2020, Section 3) for details. These methods usually infer fidelity usefulness from statistical dependence with the target, which is effective when fidelities are aligned but can increase cumulative evaluation cost over SFBO when sources are unreliable. rMFBO (Mikkola et al., 2023) mitigates this with a theoretical performance guarantee relative to SFBO even with unreliable sources, but only for a single-objective. RESCUE addresses this issue by introducing causal structure into the surrogate.

Despite extensive work on MFBO and causal BO (Aglietti et al., 2020; 2023), the literature on integrating causal inference in MF setting is limited. MO-CBO (Bhatija et al., 2025) extends causal BO to the multi-objective setting but is limited to a single fidelity, while MFACBO (Zeitler, 2025) outlines an integration of causal abstraction with MFBO but is limited to a single objective and a single discrete lower fidelity. Prior work has also shown that causal structure can improve the optimization and debugging of configurable systems by suppressing spurious correlations and transferring causally relevant information across environments (Iqbal et al., 2023; Hossen et al., 2025). RESCUE differs from these efforts by targeting multi-objective MF optimization, where the optimizer must select both an input and an evaluation fidelity while optimizing the target-fidelity Pareto front under constraints.

3 Preliminaries

3.1 Problem Setup

Consider a highly configurable system with x_i indicating the ith configuration option, which takes values $x_i \in \mathrm{Dom}(\mathrm{x}_i)$. The configuration space is the Cartesian product of all parameter domains, $\mathbb{X} = \times_{i\in[d]} \mathrm{Dom}(\mathrm{x}_i)$, where d is the number of configuration options. A configuration $\boldsymbol{x} \in \mathbb{X}$ is instantiated by setting a specific value for each option within its domain, $\boldsymbol{x} = \langle \mathrm{x}_1 = x_1, \mathrm{x}_2 = x_2, \ldots, \mathrm{x}_d = x_d \rangle$. Let $\mathbb{S}$ denote the fidelity space, and let $s_\diamond \in \mathbb{S}$ denote the target fidelity (e.g., a real system). For any fidelity $s \in \mathbb{S}$, we represent the performance objective as a black-box mapping from the joint configuration-fidelity space to M real-valued objectives, $\boldsymbol{f} : \mathbb{X} \times \mathbb{S} \to \mathbb{R}^M$, where $\boldsymbol{f} = \{f_m\}_{m=1}^{M\geq 2}$. In practice, we learn $\boldsymbol{f}$ by sampling from the configuration-fidelity space and collecting observations $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}, s) + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon}$ is a realization of $\epsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{I})$. Evaluations

at fidelities $s \neq s_\diamond$ are regarded as auxiliary sources that approximate the target ($s_\diamond$) with variable bias $\boldsymbol{\delta}(\boldsymbol{x}, s) = \boldsymbol{f}(\boldsymbol{x}, s) - \boldsymbol{f}(\boldsymbol{x}, s_\diamond)$. We consider $\boldsymbol{f}(\boldsymbol{x}, s_\diamond)$ subject to Q inequality constraints $\{h_q(\boldsymbol{x}, s)\}_{q=1}^{Q}$, which are observable across all fidelities and satisfied when $h_q(\boldsymbol{x}, s) \geq 0$. Let $c(\boldsymbol{x}, s)$ be the evaluation cost, and assume $c(\boldsymbol{x}, s) < c(\boldsymbol{x}, s_\diamond)$ for $s \neq s_\diamond$, where s could correspond to simulation tools or simplified settings. Furthermore, $\boldsymbol{f}(\boldsymbol{x}, s_\diamond)$ can be observed directly without bias, although noise may influence it. Our goal is to find the configurations $\mathbb{X}^*$ *at the target fidelity* $s_\diamond$ that yield Pareto-optimal performance while satisfying the constraints under a budget Λ, requiring the cumulative cost $C \coloneqq \sum_{t=1}^{T} c(\boldsymbol{x}_t, s_t)$ to satisfy $C \leq \Lambda$, with T denoting total interventions. Formally, with $\prec$ denoting Pareto dominance,

$$\mathbb{X}^* = \{\boldsymbol{x} \in \bar{\mathbb{X}} : \nexists\, \boldsymbol{x}' \in \bar{\mathbb{X}},\ \boldsymbol{f}(\boldsymbol{x}', s_\diamond) \prec \boldsymbol{f}(\boldsymbol{x}, s_\diamond)\}, \quad \bar{\mathbb{X}} = \{\boldsymbol{x} \in \mathbb{X} : h_q(\boldsymbol{x}, s_\diamond) \geq 0\ \forall q \in [Q]\}. \tag{1}$$

3.2 Background

Let $\mathcal{G} = (\mathbb{V}, \mathbb{A})$ denote a directed acyclic graph (DAG), where $\mathbb{V}$ is the set of observable variables and $\mathbb{A}$ is the set of directed edges capturing causal dependencies.A causal performance model (CPM) (Iqbal et al., 2022; 2023) is a structural causal model (SCM) (Pearl, 2009; Peters et al., 2017) specialized to configurable systems, where configuration options influence outputs through key performance indicators. The CPM is defined as $\mathcal{C} = \langle \mathbb{U}, \mathbb{V}, \mathbb{F} \rangle$, where $\mathbb{U}$ is a set of mutually independent exogenous variables determined by factors outside the model, and $\mathbb{F}$ is a set of deterministic functions such that $\mathrm{v}_i = f_i(Pa_{\mathcal{G}}(\mathrm{v}_i), \mathrm{u}_i)$ for each $\mathrm{v}_i \in \mathbb{V}$, with $Pa_{\mathcal{G}}(\mathrm{v}_i) \subseteq \mathbb{V} \setminus \{\mathrm{v}_i\}$ denoting the set of parents of v_i and $\mathrm{u}_i \in \mathbb{U}$. We draw a directed edge from v_i to v_j if $\mathrm{v}_i \in Pa_{\mathcal{G}}(\mathrm{v}_j)$. We partition $\mathbb{V}$ into configuration options $\mathbf{x}$, key performance indicators $\mathbf{z}$ (non-manipulable variables), outputs, comprising the performance objectives $\mathbf{y}$ and the constraint variables $\mathbf{h}$, and the fidelity s. We assume the system can be described by a CPM.

Using the rules of *do*-calculus (Pearl, 2009), a causal intervention (CI) (Peters et al., 2017) modifies $\mathcal{C}$ by replacing $\boldsymbol{f}_{\mathbf{x}}(Pa_{\mathcal{G}}(\mathbf{x}), \mathbf{u}_{\mathbf{x}})$ with $\boldsymbol{x}$, denoted by $\mathrm{do}(\mathbf{x} = \boldsymbol{x})$ which represents an intervention on $\mathbf{x}$ by setting its value to $\boldsymbol{x}$. Given $\mathcal{G}$, we can write the observational distribution as $p((\mathbf{y}, \mathbf{h}) \mid \mathbf{x} = \boldsymbol{x}, s)$ using the Causal Markov Condition (Pearl, 2009). Evaluating $p((\mathbf{y}, \mathbf{h}) \mid \mathbf{x} = \boldsymbol{x}, s)$ requires *observing* while $p((\mathbf{y}, \mathbf{h}) \mid \mathrm{do}((\mathbf{x}, \mathrm{s}) = (\boldsymbol{x}, s)))$ requires *interventions*. We intervene on the fidelity s together with $\mathbf{x}$ since the optimizer selects both. When $p((\mathbf{y}, \mathbf{h}) \mid \mathrm{do}((\mathbf{x}, \mathrm{s}) = (\boldsymbol{x}, s)))$ is identifiable from $\mathcal{G}$ (Galles & Pearl, 1995), the rules of *do*-calculus reduce it to adjustment formulas over observed variables. The adjustment formulas involve integrals over the adjustment sets, which we approximate with Monte Carlo (MC) estimates from observational data[1] $\widehat{\mathcal{D}}$, yielding $\widehat{p}((\mathbf{y}, \mathbf{h}) \mid \mathrm{do}((\mathbf{x}, \mathrm{s}) = (\boldsymbol{x}, s))) \approx p((\mathbf{y}, \mathbf{h}) \mid \mathrm{do}((\mathbf{x}, \mathrm{s}) = (\boldsymbol{x}, s)))$. For notational simplicity, we denote $\boldsymbol{f}_{\mathrm{do}}(\boldsymbol{x}, s) \coloneqq \mathbb{E}_{\widehat{p}((\mathbf{y}, \mathbf{h}) \mid \mathrm{do}((\mathbf{x}, \mathrm{s}) = (\boldsymbol{x}, s)))}[(\mathbf{y}, \mathbf{h})] \in \mathbb{R}^J$ with $J = M + Q$.

In this paper, we learn $\mathcal{G}$ from $\widehat{\mathcal{D}}$ using the PC algorithm (Spirtes et al., 2001) and DirectLiNGAM (Shimizu et al., 2011). Both algorithms assume *causal sufficiency*, the absence of unmeasured common causes of observed variables. RESCUE takes the CPM as an input, so practitioners can use causal discovery methods that relax this assumption, such as FCI-type algorithms (Spirtes et al., 2001). Section 5 quantifies the effect of causal-model error on the optimization guarantee. We enforce domain-specific structural constraints, such as disallowing directed edges from $\mathbf{y}$ to $\mathbf{x}$, to facilitate learning $\mathcal{G}$ with limited data (see Hossen et al. (2023, Section 3) for details). RESCUE assumes an identifiable causal graph, adequate $\widehat{\mathcal{D}}$ and valid causal estimates.

4 Methodology

In this section, we introduce a joint probabilistic surrogate for the objectives and constraints of Eq. (1) with a CPM-derived prior. We then design causal acquisition strategy balancing information and cost across fidelities while accounting for feasibility. Finally, we present the RESCUE algorithm to find optimal configurations under a finite evaluation budget.

[1]We assume access to either observational data $\widehat{\mathcal{D}}$ (e.g., system logs, which are readily available in most systems) or a given DAG.

4.1 MULTI-OUTPUT MULTI-FIDELITY CAUSAL GP

Let $\mathcal{D} = \{(\boldsymbol{x}_i, s_i, \boldsymbol{y}_i)\}_{i=1}^N$ be a finite collection of N interventional samples. We write $\bar{\boldsymbol{f}} = (\boldsymbol{f}, h_1, \ldots, h_Q)$ for the objectives and constraints stacked into $J = M+Q$ outputs, and $\boldsymbol{y} \in \mathbb{R}^J$ for its noisy observation, i.e., $\boldsymbol{y}_i = \bar{\boldsymbol{f}}(\boldsymbol{x}_i, s_i) + \boldsymbol{\epsilon}_i$, where $\boldsymbol{\epsilon}_i$ is a realization of $\epsilon_i \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$ i.i.d.[2]. The likelihood of $\mathcal{D}$ is $p(\mathcal{D} \mid \bar{\boldsymbol{f}}, \sigma^2) = \prod_{i=1}^N \mathcal{N}\big(\boldsymbol{y}_i \,;\, \bar{\boldsymbol{f}}(\boldsymbol{x}_i, s_i), \sigma^2 \boldsymbol{I}\big)$.

We place a causal prior (Aglietti et al., 2020) on $\bar{\boldsymbol{f}}(\boldsymbol{x}, s)$. The causal prior is computed to estimate the expected interventional outcome of a configuration on the objectives. We construct the prior mean and variance of a GP using the observational data $\widehat{\mathcal{D}}$ and the corresponding CPM $\mathcal{C}$. The MF-CGP is a multi-output GP over J outputs and $j \in [J]$. For notational brevity[3], we define $\boldsymbol{\chi} := (\boldsymbol{x}, s)$ and $\boldsymbol{\chi}' := (\boldsymbol{x}', s')$. The MF-CGP is defined as

$$\bar{\boldsymbol{f}}(\boldsymbol{\chi}) \sim \mathrm{GP}\left(\boldsymbol{\mu}(\boldsymbol{\chi}), k_{j,j'}(\boldsymbol{\chi}, \boldsymbol{\chi}')\right) \text{ with } \boldsymbol{\mu}(\boldsymbol{\chi}) = \bar{\boldsymbol{f}}_{\mathrm{do}}(\boldsymbol{\chi}) \text{ and}$$
$$k_{j,j'}\left(\boldsymbol{\chi}, \boldsymbol{\chi}'\right) = \left(k_{\mathrm{in}}(\boldsymbol{x}, \boldsymbol{x}')\, k_{\mathrm{fid}}(s, s') + \langle \widehat{\boldsymbol{\sigma}}(\boldsymbol{\chi}), \widehat{\boldsymbol{\sigma}}(\boldsymbol{\chi}')\rangle\right) \otimes k_{\mathrm{obj}}(j, j'),$$

where $\widehat{\boldsymbol{\sigma}}(\boldsymbol{\chi}) = \left(\mathrm{Var}_{\widehat{p}((\mathbf{y},\mathbf{h})|\mathrm{do}((\mathbf{x},\mathbf{s})=(\boldsymbol{x},s)))}[(\mathbf{y}, \mathbf{h})]\right)^{1/2}$ with Var the componentwise variance estimated from $\widehat{\mathcal{D}}$, and $\langle \widehat{\boldsymbol{\sigma}}(\boldsymbol{\chi}), \widehat{\boldsymbol{\sigma}}(\boldsymbol{\chi}')\rangle = \widehat{\boldsymbol{\sigma}}(\boldsymbol{\chi})^\top \widehat{\boldsymbol{\sigma}}(\boldsymbol{\chi}')$ is a finite-rank covariance encoding the uncertainty of the causal estimates. The kernel $k_{j,j'}(\boldsymbol{\chi}, \boldsymbol{\chi}') = \mathrm{Cov}(\bar{f}_j(\boldsymbol{\chi}), \bar{f}_{j'}(\boldsymbol{\chi}'))$ captures spatial similarity and inter-output relationships. Here, $k_{\mathrm{in}}(\cdot, \cdot)$ captures covariance between configurations, $k_{\mathrm{fid}}(\cdot, \cdot)$ captures similarity across fidelities, while the cross-fidelity bias $\boldsymbol{\delta}(\boldsymbol{x}, s)$ is captured by $\bar{\boldsymbol{f}}_{\mathrm{do}}(\boldsymbol{x}, s)$, and $k_{\mathrm{obj}}(\cdot, \cdot)$ is an index kernel over the outputs, defined as $k_{\mathrm{obj}}(j, j') = (\boldsymbol{L}\boldsymbol{L}^\top + \mathrm{diag}(\boldsymbol{v}))_{j,j'}$, where $\boldsymbol{L} \in \mathbb{R}^{J \times J}$ and $\boldsymbol{v} \in \mathbb{R}^J_{\geq 0}$ are learnable, so that the output covariance matrix is positive semi-definite and captures correlations among the outputs. For both k_{in} and k_{fid} we use the radial basis function (RBF) kernel, $k_{\mathrm{in}}(\boldsymbol{x}, \boldsymbol{x}') = \exp(-\frac{\|\boldsymbol{x}-\boldsymbol{x}'\|^2}{2l_{\mathrm{in}}^2})$ and $k_{\mathrm{fid}}(s, s') = \exp(-\frac{(s-s')^2}{2l_{\mathrm{fid}}^2})$. The joint posterior over objectives and constraints allows low-fidelity constraint observations to inform target-fidelity feasibility through the cross-fidelity kernel.

Given this GP prior, the posterior distribution over $\bar{\boldsymbol{f}}(\boldsymbol{\chi})$ can be derived analytically via standard GP updates. For notational brevity, we write $\boldsymbol{\chi}_i := (\boldsymbol{x}_i, s_i)$ and $\boldsymbol{\chi}_{1:N} := (\boldsymbol{\chi}_1, \ldots, \boldsymbol{\chi}_N)$. The posterior distribution is given as, $p(\bar{\boldsymbol{f}}(\boldsymbol{\chi})|\mathcal{D}) = \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\chi}|\mathcal{D}),\, \boldsymbol{K}(\boldsymbol{\chi}, \boldsymbol{\chi}|\mathcal{D}))$, where $\boldsymbol{\mu}(\boldsymbol{\chi}|\mathcal{D}) = \boldsymbol{\mu}(\boldsymbol{\chi}) + \boldsymbol{K}_*(\boldsymbol{\chi}, \boldsymbol{\chi}_{1:N})^\top (\boldsymbol{\Sigma} + \sigma^2 \boldsymbol{I})^{-1} (\boldsymbol{y}_{1:N} - \boldsymbol{\mu}(\boldsymbol{\chi}_{1:N}))$ and $\boldsymbol{K}(\boldsymbol{\chi}, \boldsymbol{\chi}'|\mathcal{D}) = k(\boldsymbol{\chi}, \boldsymbol{\chi}') - \boldsymbol{K}_*(\boldsymbol{\chi}, \boldsymbol{\chi}_{1:N})^\top (\boldsymbol{\Sigma} + \sigma^2 \boldsymbol{I})^{-1} \boldsymbol{K}_*(\boldsymbol{\chi}', \boldsymbol{\chi}_{1:N})$ with $\boldsymbol{K}_*(\boldsymbol{\chi}, \boldsymbol{\chi}_{1:N}) = [k(\boldsymbol{\chi}, \boldsymbol{\chi}_1) \;\ldots\; k(\boldsymbol{\chi}, \boldsymbol{\chi}_N)]^\top$ representing an $(NJ) \times J$ matrix with each $k(\boldsymbol{\chi}, \boldsymbol{\chi}_i) = [k_{j,j'}(\boldsymbol{\chi}, \boldsymbol{\chi}_i)]_{j,j'=1}^J$ representing a $J \times J$ covariance matrix, and $\boldsymbol{\Sigma} = [k(\boldsymbol{\chi}_i, \boldsymbol{\chi}_{i'})]_{i,i'=1}^N$ is the $(NJ) \times (NJ)$ Gram matrix.

4.2 VALUE-OF-CAUSAL-INFORMATION HYPERVOLUME KNOWLEDGE GRADIENT

We extend the HVKG approach (Daulton et al., 2023) by integrating causal information into the acquisition to infer a hypervolume-maximizing *feasible* Pareto set at $s_\diamond$ for the constrained problem in Eq. (1). The optimization algorithm proceeds in rounds (hereafter referred to as iterations). At each iteration t, the acquisition function selects a configuration–fidelity pair $(\boldsymbol{x}, s)$ for intervention.

Definition 4.1. Given a finite Pareto frontier $\mathbb{P} \subset \mathbb{R}^M$ and a reference point $\boldsymbol{r} \in \mathbb{R}^M$, the hypervolume (HV) is the volume of the dominated space[4]: $\mathrm{HV}(\mathbb{P}, \boldsymbol{r}) = \lambda(\bigcup_{\boldsymbol{y}' \in \mathbb{P}} \{\boldsymbol{y} \in \mathbb{R}^M : \boldsymbol{y}' \preceq \boldsymbol{y} \preceq \boldsymbol{r}\})$, where λ is the Lebesgue measure.

Let $\mathcal{D}_{(\boldsymbol{x},s)} := \mathcal{D} \cup \{(\boldsymbol{x}, s, \boldsymbol{y})\}$ denote the dataset augmented with an additional observation $(\boldsymbol{x}, s, \boldsymbol{y})$. The causal prior is learned from observational data, and the acquisition maintains an estimate of how much this prior should be trusted. To represent this trust, let $\boldsymbol{\omega} \in [0, 1]^J$ be a *per-output reliability*. Let $\mathcal{M}_{\boldsymbol{\omega}}$ denote the MF-CGP whose causal prior contribution is scaled by $\boldsymbol{\omega}$, with prior mean $\boldsymbol{\omega} \odot \bar{\boldsymbol{f}}_{\mathrm{do}} + (\boldsymbol{1} - \boldsymbol{\omega}) \odot \boldsymbol{c}$, where $\boldsymbol{c}$ is a learned constant mean, and $\langle \widehat{\boldsymbol{\sigma}}, \widehat{\boldsymbol{\sigma}}' \rangle$ replaced by $\langle \boldsymbol{\omega} \odot \widehat{\boldsymbol{\sigma}}, \boldsymbol{\omega} \odot \widehat{\boldsymbol{\sigma}}' \rangle$. Thus, $\boldsymbol{\omega} = \boldsymbol{1}$ uses the full CPM contribution, whereas $\boldsymbol{\omega} = \boldsymbol{0}$ removes it entirely. The kernel hyperparameters $\boldsymbol{\theta}$ of $\mathcal{M}_{\boldsymbol{0}}$ and $\mathcal{M}_{\boldsymbol{1}}$ are fit by maximizing the log marginal likelihood under the priors

[2] To avoid clutter, $\boldsymbol{\epsilon}_i$ is omitted in the remainder of the paper.

[3] In this paper, $(\boldsymbol{x}, s)$ and $\boldsymbol{\chi}$ are used interchangeably.

[4] For brevity, $\boldsymbol{r}$ is omitted from the remainder of the paper.

$p(\boldsymbol{\theta})$. The causal trust $\widehat{\boldsymbol{\omega}}$ is a model weight computed from the leave-one-out (LOO) predictive log likelihood (Sundararajan & Keerthi, 2001; Yao et al., 2018):

$$\widehat{\omega}_j \;=\; \frac{\prod_i p\big(y_{i,j} \mid \mathcal{D}_{-i}, \mathcal{M}_{\mathbf{1}}\big)}{\prod_i p\big(y_{i,j} \mid \mathcal{D}_{-i}, \mathcal{M}_{\mathbf{0}}\big) + \prod_i p\big(y_{i,j} \mid \mathcal{D}_{-i}, \mathcal{M}_{\mathbf{1}}\big)}, \tag{2}$$

where $\mathcal{D}_{-i}$ omits the i-th observation. The posterior predictive of each output is the mixture $p(\bar{f}_j \mid \mathcal{D}) = \widehat{\omega}_j\, p(\bar{f}_j \mid \mathcal{D}, \mathcal{M}_{\mathbf{1}}) + (1-\widehat{\omega}_j)\, p(\bar{f}_j \mid \mathcal{D}, \mathcal{M}_{\mathbf{0}})$ with posterior mean $\boldsymbol{\mu}(\boldsymbol{\chi} \mid \mathcal{D}, \widehat{\boldsymbol{\omega}})$.

We follow Daulton et al. (2023, Section 4) to approximate a finite Pareto set of at most $n_{\mathbb{P}}$ configurations. For a candidate set $\mathbb{B} \subseteq \mathbb{X}$ with $|\mathbb{B}| \le n_{\mathbb{P}}$, let $\mathbb{Y}(\mathbb{B} \mid \mathcal{D}) := \{\boldsymbol{\mu}_{1:M}(\boldsymbol{x}, s_\diamond \mid \mathcal{D}, \widehat{\boldsymbol{\omega}}) : \boldsymbol{x} \in \mathbb{B},\ \mu_{M+q}(\boldsymbol{x}, s_\diamond \mid \mathcal{D}, \widehat{\boldsymbol{\omega}}) \ge 0\ \forall q \in [Q]\}$ denote the posterior mean objective values of its members whose posterior mean constraints are satisfied, following the constrained hypervolume formulation of Daulton et al. (2020). The value of a dataset $\mathcal{D}$ is

$$\nu(\mathcal{D}) := \max_{\mathbb{B}\subseteq\mathbb{X},\ |\mathbb{B}|\le n_{\mathbb{P}}} \mathrm{HV}\big[\mathbb{Y}(\mathbb{B} \mid \mathcal{D})\big]. \tag{3}$$

The Value-of-Causal-Information Hypervolume Knowledge Gradient (VoCI-HVKG) is the cost-normalized expected improvement of $\nu(\mathcal{D})$:

$$\alpha(\boldsymbol{x}, s) = \frac{1}{c(\boldsymbol{x}, s)}\, \mathbb{E}_{(\mathbf{y},\mathbf{h})\sim p(\cdot \mid \boldsymbol{x}, s, \mathcal{D})} \left[\nu\big(\mathcal{D}_{(\boldsymbol{x},s)}\big) - \nu(\mathcal{D})\right],$$

where $p(\cdot \mid \boldsymbol{x}, s, \mathcal{D})$ is the mixture posterior predictive distribution, and $\mathcal{D}_{(\boldsymbol{x},s)}$ updates both candidate posteriors and the weights $\widehat{\boldsymbol{\omega}}$ via Eq. (2). Conceptually, VoCI-HVKG quantifies the expected increase in the feasible hypervolume at the target fidelity by combining the predictive uncertainty of the MF-CGP with the *value of causal information* carried by the prior, weighted by the causal trust $\widehat{\boldsymbol{\omega}}$. Note that the query $(\boldsymbol{x}, s)$ is not restricted to feasible configurations, because an infeasible configuration can still be informative about the feasible optimum. At each iteration t, the next query is

$$(\boldsymbol{x}_t, s_t) = \operatorname*{arg\,max}_{(\boldsymbol{x},s)\in\mathbb{X}\times\mathbb{S}} \alpha(\boldsymbol{x}, s). \tag{4}$$

Although VoCI-HVKG cannot be computed analytically, following the one-shot formulation of Daulton et al. (2023, Theorem B.1) we obtain a consistent estimator by approximating the outer expectation via Monte Carlo:

$$\widehat{\alpha}(\boldsymbol{x}, s) = \frac{1}{c(\boldsymbol{x}, s)} \left[\frac{1}{N_{\mathrm{f}}} \sum_{i=1}^{N_{\mathrm{f}}} \nu\big(\mathcal{D}^i_{(\boldsymbol{x},s)}\big) - \nu(\mathcal{D})\right],$$

where $\mathcal{D}^i_{(\boldsymbol{x},s)} = \mathcal{D} \cup \{(\boldsymbol{x}, s, \boldsymbol{y}^i)\}$ and each $\boldsymbol{y}^i$ is a "fantasy" sample from the mixture posterior, output j drawn from $\mathcal{M}_{\mathbf{1}}$ with probability $\widehat{\omega}_j$. We condition both candidates on each fantasy sample. We approximate the constraint indicators in ν with differentiable sigmoids of the posterior mean constraints (Daulton et al., 2020).

Algorithm 1 RESCUE Algorithm

Require: $\widehat{\mathcal{D}}$, $\mathbb{X}$, $\mathbb{S}$, $s_\diamond$, $c(\cdot,\cdot)$, budget Λ
Ensure: Recommended Pareto set
1: Initialize $\mathcal{D}_0 = \{(\boldsymbol{x}_i, s_i, \boldsymbol{y}_i)\}_{i=1}^{N_0}$ (Appendix A.1), $t = 0$, and $C_0 = 0$
2: Construct the CPM: $\mathcal{C}(\mathcal{G}|\widehat{\mathcal{D}})$
3: **while** $C_t \le \Lambda$ **do**
4: $\quad t \leftarrow t + 1$
5: $\quad \mathcal{M}(\boldsymbol{x}, s \mid \mathcal{D}_{0:t-1}, \mathcal{C}) \leftarrow$ Train a MF-CGP
6: $\quad$ Find next $(\boldsymbol{x}_t, s_t)$ by optimizing Eq. (4)
7: $\quad$ **if** $C_{t-1} + c(\boldsymbol{x}_t, s_t) > \Lambda$ **then**
8: $\quad\quad$ **break**
9: $\quad$ **end if**
10: $\quad$ Obtain $\boldsymbol{y}_t$ by intervening $\boldsymbol{x}_t$ at s_t
11: $\quad$ Augment $\mathcal{D}_{0:t} \leftarrow \mathcal{D}_{0:t-1} \cup \{(\boldsymbol{x}_t, s_t, \boldsymbol{y}_t)\}$
12: $\quad$ Update total cost $C_t \leftarrow C_{t-1} + c(\boldsymbol{x}_t, s_t)$
13: **end while**
14: Obtain Pareto front of the posterior mean at $s_\diamond$ subject to $\mathbb{E}[h_q(\boldsymbol{x}, s_\diamond) \mid \mathcal{D}] \ge 0\ \forall q \in [Q]$

We present the complete RESCUE algorithm in Algorithm 1. To initialize $\mathcal{D}_0$, we use a cost-aware sampling strategy described in Appendix A.1 (Algorithm 2). The final recommendation extracts the Pareto front of the posterior mean at the target fidelity (Line 14), a separation between query selection and recommendation that is standard in BO (Daulton et al., 2023). The time complexity of Algorithm 1 has three main components: (i) initial CPM construction on $\mathbb{V}$ from $\widehat{\mathcal{D}}$ using PC or DirectLiNGAM, which requires $\mathcal{O}(|\widehat{\mathcal{D}}|\,|\mathbb{V}|^i)$ for PC, where i is the maximum degree of $\mathbb{V}$ in $\mathcal{G}$, or $\mathcal{O}(|\widehat{\mathcal{D}}|\,|\mathbb{V}|^3 + |\mathbb{V}|^4)$ for DirectLiNGAM; (ii) causal inference via do-calculus for computing $\bar{\boldsymbol{f}}_{\mathrm{do}}(\boldsymbol{x}, s)$, with complexity $\mathcal{O}(|\mathbb{V}|^2 + |\widehat{\mathcal{D}}|)$ per query; and (iii) MF-CGP inference per iteration, with complexity $\mathcal{O}((NJ)^3)$ to construct and invert the Gram matrix. The space complexity is $\mathcal{O}((NJ)^2)$ for storing the Gram matrix, where N is the number of observations.

5 THEORETICAL ANALYSIS

We analyze RESCUE's performance when auxiliary sources poorly approximate the target fidelity, extending the single-objective regret bounds of Mikkola et al. (2023) to the constrained multi-objective setting. We establish an *anytime guarantee* that bounds the target-fidelity hypervolume regret of the recommended Pareto set at every iteration. We also establish a *robustness guarantee* showing that, under Assumption 5.2, this bound is independent of the magnitude of the cross-fidelity bias $\boldsymbol{\delta}(\boldsymbol{x}, s)$. For a variant of Algorithm 1 with a periodic target-fidelity exploration step, we additionally establish a convergence guarantee with a sublinear rate (Appendix B.6). We provide proofs in Appendix B.

Throughout, we view the MF-CGP as a scalar GP over configuration–fidelity–output triples, with kernel $\tilde{k}\big((\boldsymbol{\chi}, j), (\boldsymbol{\chi}', j')\big) = k_{j,j'}(\boldsymbol{\chi}, \boldsymbol{\chi}')$, posterior mean $\boldsymbol{\mu}_t$, and standard deviation $\boldsymbol{\sigma}_t$ after t interventions. We write $w((\boldsymbol{\chi}, j)) = w_j(\boldsymbol{\chi})$ for the j-th component of $\boldsymbol{w}$, and similarly for $\bar{\boldsymbol{f}}_{\mathrm{do}}$. We write γ_t for the maximum information gain after t (vector-valued) interventions and $\bar{\mathbb{X}}$ for the feasible region of Eq. (1). As is standard in GP-based BO analyses, we fix the kernel hyperparameters for the analysis. In practice, we select them by maximizing the log marginal likelihood.

Assumption 5.1 (GP regularity)**.** The kernel $\tilde{k}$ is bounded, $\tilde{k} \leq 1$, and observations are corrupted by the independent Gaussian noise of Eq. (1) with variance σ^2.

Assumption 5.2 (Causal prior quality)**.** There exists w in the RKHS $\mathbb{H}_{\tilde{k}}$ of $\tilde{k}$ with $\|w\|_{\mathbb{H}_{\tilde{k}}} \leq B$ such that $\sup_{(\boldsymbol{x},s)} \big\| \bar{\boldsymbol{f}}(\boldsymbol{x}, s) - \bar{\boldsymbol{f}}_{\mathrm{do}}(\boldsymbol{x}, s) - \boldsymbol{w}(\boldsymbol{x}, s) \big\|_\infty \leq \xi$. We refer to $\xi = 0$ as the *well-specified* setting and to $\xi > 0$ as the *misspecified* setting.

Assumption 5.3 (Bounded objectives)**.** At the target fidelity, all objective values lie in an axis-aligned box of side length R whose upper corner is the reference point $\boldsymbol{r}$.

Assumption 5.4 (Cost structure)**.** Costs are normalized so that $c(\boldsymbol{x}, s_\diamond) = 1$ and $c_{\min} \leq c(\boldsymbol{x}, s) \leq 1$ for some $c_{\min} > 0$.

Assumption 5.5 (Constraint margin regularity)**.** There exists $L' \geq 0$ such that $\mathrm{HV}^* - \mathrm{HV}^*_\tau \leq L'\tau$ for all $\tau \geq 0$, where HV^*_τ is the optimal hypervolume over $\bar{\mathbb{X}}_\tau := \{\boldsymbol{x} : h_q(\boldsymbol{x}, s_\diamond) \geq \tau \ \forall q\}$.

Lemma 5.6 (Hypervolume stability)**.** *Let $\mathbb{X}_0 \subseteq \mathbb{X}$. Suppose $\boldsymbol{a}, \boldsymbol{b} : \mathbb{X}_0 \to \mathbb{R}^M$ have outputs in the box of Assumption 5.3 with $\sup_{\boldsymbol{x} \in \mathbb{X}_0} \|\boldsymbol{a}(\boldsymbol{x}) - \boldsymbol{b}(\boldsymbol{x})\|_\infty \leq \epsilon$ for a perturbation tolerance $\epsilon \geq 0$. Then $\big|\mathrm{HV}[\{\boldsymbol{a}(\boldsymbol{x})\}_{\boldsymbol{x} \in \mathbb{X}_0}] - \mathrm{HV}[\{\boldsymbol{b}(\boldsymbol{x})\}_{\boldsymbol{x} \in \mathbb{X}_0}]\big| \leq L\,\epsilon$, where $L := 2MR^{M-1}$.*

Lemma 5.7 (Confidence bound)**.** *Under Assumptions 5.1–5.2, for any $\rho \in (0, 1)$, with probability at least $1 - \rho$, for all $\boldsymbol{x}$, s, $t \geq 1$, and all outputs j,*

$$\big|\bar{f}_j(\boldsymbol{x}, s) - \mu_{t,j}(\boldsymbol{x}, s)\big| \leq \beta_t^{1/2} \sigma_{t,j}(\boldsymbol{x}, s) + \xi\Big(1 + \tfrac{1}{\sigma}\sqrt{Jt}\ \sigma_{t,j}(\boldsymbol{x}, s)\Big),$$

where $\beta_t^{1/2} = B + \sqrt{2\big(\gamma_t + 1 + \ln(1/\rho)\big)}$.

For the analysis, $\widehat{\mathbb{P}}_T$ denotes the Pareto set of the target-fidelity objective posterior means $(\mu_{T,1}, \ldots, \mu_{T,M})$ after T interventions over the estimated feasible region $\widehat{\mathbb{X}}_T := \{\boldsymbol{x} : \mu_{T,M+q}(\boldsymbol{x}, s_\diamond) \geq \beta_T^{1/2} \sigma_{T,M+q}(\boldsymbol{x}, s_\diamond)\ \forall q\}$. This is stricter than the posterior-mean feasibility of Line 14 in Algorithm 1. The objective posterior means are projected onto the box of Assumption 5.3.

Theorem 5.8 (Anytime guarantee for RESCUE)**.** *Let Assumptions 5.1–5.5 hold and fix the kernel hyperparameters. For any $\rho \in (0, 1)$, with probability at least $1 - \rho$, for all $T \geq 1$,*

$$\mathrm{HV}^* - \mathrm{HV}\big[\{\boldsymbol{f}(\boldsymbol{x}, s_\diamond)\}_{\boldsymbol{x} \in \widehat{\mathbb{P}}_T}\big] \ \leq\ 2\big(L + L'\big)\left(\beta_T^{1/2}\,\bar{\sigma}_T \ +\ \xi\Big(1 + \tfrac{1}{\sigma}\sqrt{JT}\,\bar{\sigma}_T\Big)\right), \tag{5}$$

where $\bar{\sigma}_T := \sup_{\boldsymbol{x} \in \mathbb{X}} \|\boldsymbol{\sigma}_T(\boldsymbol{x}, s_\diamond)\|_2$ is the realized aggregated posterior uncertainty at the target fidelity. Every recommended configuration $\boldsymbol{x} \in \widehat{\mathbb{P}}_T$ satisfies $h_q(\boldsymbol{x}, s_\diamond) \geq -\xi\big(1 + \frac{1}{\sigma}\sqrt{JT}\,\bar{\sigma}_T\big)$ for all q.

Corollary 5.9 (Unreliable auxiliary source robustness)**.** *Suppose Assumptions 5.1–5.5 hold, and let $\boldsymbol{\delta}(\boldsymbol{x}, s) := \boldsymbol{f}(\boldsymbol{x}, s) - \boldsymbol{f}(\boldsymbol{x}, s_\diamond)$ denote the cross-fidelity bias of auxiliary sources at fidelity $s \neq s_\diamond$.*

(i) For any horizon T and any $\rho \in (0, 1)$, the bound of Theorem 5.8 depends on the auxiliary sources only through the constants (B, ξ) of Assumption 5.2 and the realized uncertainty $\bar{\sigma}_T$,



Figure 1: Real robot with obstacles (left), without obstacles (middle), and *Gazebo* simulation (right).

and not on the magnitude of the cross-fidelity bias $\boldsymbol{\delta}$. In particular, for any family of auxiliary sources satisfying Assumption 5.2 with common (B, ξ), the bound holds with the same constants as $\sup_{\boldsymbol{x}, s \neq s_\diamond} \|\boldsymbol{\delta}(\boldsymbol{x}, s)\| \to \infty$.

(ii) If the recommendation is computed from the posterior conditioned on the target-fidelity interventions alone, Theorem 5.8 holds for such a recommendation with T replaced by the number of target-fidelity interventions, and Assumption 5.2 required only at $s_\diamond$.

The mechanism behind Corollary 5.9 is structural. Because the fidelity s is a variable in the CPM, the causal prior $\bar{\boldsymbol{f}}_{\mathrm{do}}(\boldsymbol{x}, s)$ models the fidelity dependence of objectives and constraints from observational data, and the residual of Assumption 5.2 can remain small even when auxiliary sources are arbitrarily biased. A large bias can affect the queries selected by the acquisition and the rate at which $\bar{\sigma}_T$ is reduced, but the bound does not depend on its magnitude. The guarantee also extends to the reliability mixture by a union bound when Assumption 5.2 holds for $\boldsymbol{\omega} \in \{\mathbf{0}, \mathbf{1}\}$. In comparison, Mikkola et al. (2023) provide a regret-difference bound that grows exponentially with T, which the authors state is useful only in the first rounds. Their analysis does not provide a convergence rate.

6 Experiments

Our evaluation seeks to answer three Research Questions (RQs): (1) To what extent RESCUE improves sample efficiency and optimization performance under a fixed budget over the baselines? (2) How does the relative performance of RESCUE change with problem dimensionality and the presence of constraints? (3) To what extent VoCI-HVKG relies on the causal prior during optimization?

6.1 Experimental settings

We compare RESCUE against qEHVI (Daulton et al., 2020), MOMF (Irshad et al., 2024), and MF-HVKG (Daulton et al., 2023). We use BoTorch implementations (Balandat et al., 2020) for all baselines. The baselines use the multi-task GP of section 4.1 without the causal prior mean and the causal covariance term. For qEHVI, we treat s as a design variable. On the constrained problems, every method models the constraints as additional surrogate outputs and uses the feasibility-weighted form of its acquisition function (Daulton et al., 2020). See Appendix A for full experimental details[5].

Following the evaluation criterion of Daulton et al. (2023), every method obtains an approximate Pareto set by running NSGA-II (Deb et al., 2002) on the posterior mean of its *own* surrogate at the target fidelity. We then compute the *true* objectives and constraints values at these configurations at the target fidelity and the resulting HV dominated by the true Pareto frontier of the model-selected Pareto set. On constrained problems, configurations must first be feasible under the surrogate's constraint posterior and then satisfy the true constraints at the target fidelity. Only truly feasible configurations contribute to the HV. The true HV^*, computed from the ground-truth oracle, is method independent. Writing HV_t for the inferred HV at iteration t, the log HV regret is $\log_{10}(\mathrm{HV}^* - \mathrm{HV}_t)$. We also report GD^+ and IGD^+ (Ishibuchi et al., 2015) of the model-selected Pareto set against the true Pareto frontier, measuring convergence and coverage, respectively. We report means and ± 2 standard errors over 20 independent runs. All methods start from the same per-run initial dataset.

Synthetic Problems We consider three synthetic MF problems: (i) Branin-Currin (BC) ($d = 2, M = 2$) and (ii) Park ($d = 4, M = 2$) from Irshad et al. (2024), with cost $c(\boldsymbol{x}, s) = \exp(4.8s)$, and (iii) Synthetic SCM (SSCM) ($d = 4, M = 2$), an MF variant of the structural causal model

[5]Code for reproducing the experiments is available at `https://anonymous.4open.science/r/rescue-527D`

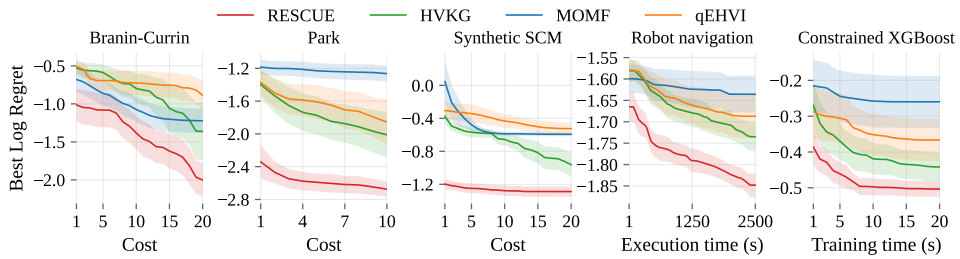


Figure 3: Cumulative intervention cost excluding the initial trials vs. best log HV regret (lower is better), where "best" denotes the minimum regret achieved up to that cost.

of Bhatija et al. (2025) (Synthetic 2), with its unit intervention costs scaled by the fidelity, $c(\boldsymbol{x}, s) = 4\,(0.1 + 0.9s)$.

Real-World Problems (i) Robot Navigation (RN) ($d = 25, M = 2, Q = 1$) from Hossen et al. (2025) involves tuning planner, controller, and costmap parameters for *Turtlebot 4* navigation in a complex environment. The target system operates with obstacles and narrow passages, while two sources consist of a *Turtlebot 4* in an obstacle-free setting and a *Gazebo* simulation (Fig. 1). The task is deemed successful if the robot reaches three goal locations and docks autonomously to a charging station. We minimize energy per meter and collision-risk score subject to a task completion rate of at least 0.95, and cost is the sum of the task completion time and a setup time of 70 seconds. For reproducibility, we curate a real-data-derived surrogate benchmark from real-robot samples and run all robot navigation experiments on it (Appendix A.7). For AutoML problems we use YAHPO Gym (Pfisterer et al., 2022) where fidelity is the fraction of the training set and cost is the training time. Specifically, (ii) Constrained XGBoost HPO (CXGB) ($d = 12, M = 2, Q = 1$) with objectives $\mathrm{logloss}$ and main effect complexity (mec) while enforcing $\mathrm{ramtrain} \leq 15.64$ MB.

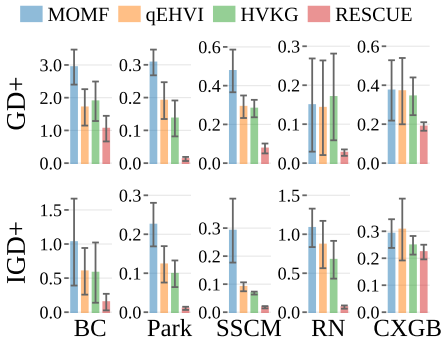


Figure 2: Convergence (top) and coverage (bottom) of Pareto set by each method.

For RESCUE, we assume the DAG is unknown and learn the CPM from $\widehat{\mathcal{D}}$, sampled uniformly at random from each problem as a stand-in for system logs. No method uses $\widehat{\mathcal{D}}$ for optimization. RESCUE uses it to learn the CPM only, and a known DAG can skip or complement causal discovery. For HVKG and VoCI-HVKG we use 8 fantasies and $n_{\mathbb{P}} = 10$ (6 and 8 for robot navigation).

6.2 Results

Fig. 3 shows that RESCUE improves sample efficiency and optimization performance over the baselines. HVKG is the strongest baseline. Fig. 2 shows that RESCUE converges closer to the true Pareto front than the baselines on all problems and covers it better on all but constrained XGBoost HPO, where it matches HVKG. On the constrained problems, RESCUE has the lowest constraint violation rate (Fig. 4), while MOMF has the highest, consistent with the acquisition function misalignment reported by Daulton et al. (2023) where fidelity is treated as an additional objective. We attribute these gains to the causal prior encoded in the MF-CGP, which captures the causal relationships between configurations, fidelities, objectives and constraints and focuses evaluations on causally promising regions of the configuration space. These results address RQ1 and RQ2 by showing that RESCUE improves sample efficiency under a fixed budget and performs consistently across low and high dimensional, constrained and unconstrained problems.

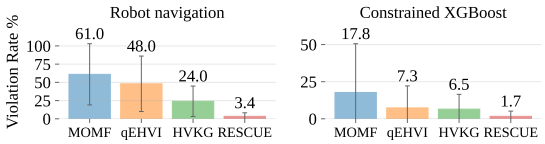


Figure 4: Fraction of iterations with no feasible configuration in the recommended Pareto set.

Ablations We investigate (i) the causal prior's contribution by setting $\boldsymbol{\omega} = \mathbf{0}$, (ii) robustness to misspecification by pairing each configuration and fidelity in $\widehat{\mathcal{D}}$ with objective and constraint

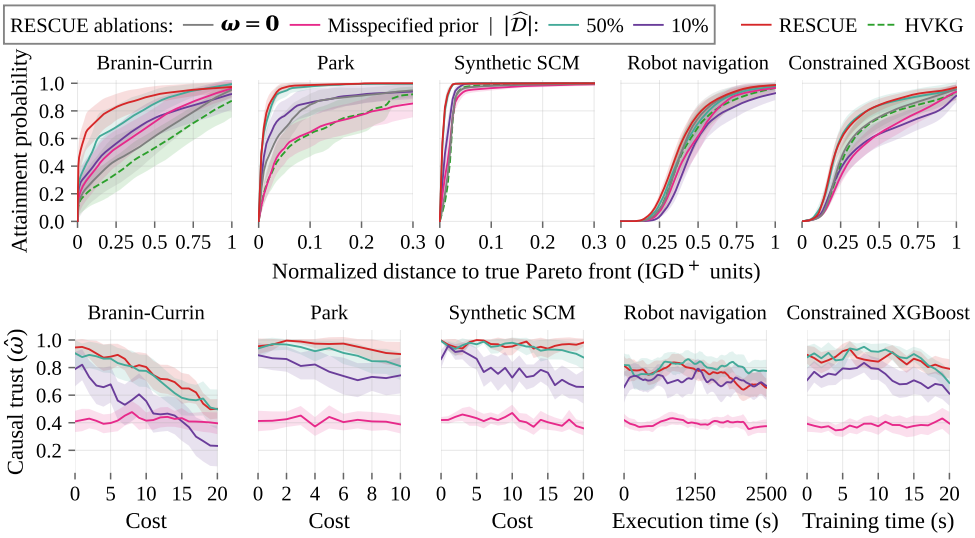


Figure 5: The attainment probability of the recommended Pareto set against the normalized distance to the true Pareto front (top) and output-averaged causal trust $\widehat{\omega}$ over optimization trajectory (bottom).

values measured at a different configuration and fidelity, inducing incorrect CPM relationships, and (iii) observational data requirements by learning the CPM from 10% and 50% of $\widehat{\mathcal{D}}$ (Appendix A.5). Fig. 5 shows that the causal prior improves RESCUE even with half of $\widehat{\mathcal{D}}$. The causal trust $\widehat{\omega}$ is lowest under misspecification and decreases as interventional data accumulate, indicating that VoCI-HVKG relies on the causal prior only to the extent it is informative, addressing RQ3. Moreover, even with a misspecified prior, RESCUE remains competitive with HVKG on all problems.

7 Discussion

To apply RESCUE to a novel problem, the practitioner must identify (i) configuration options, (ii) performance metrics, and (iii) key performance indicators. The practitioner chooses the abstraction level of CPM variables, from high-level system options to hardware.

The CPM introduces a strong inductive bias through the GP prior mean and covariance. A misspecified or overconfident prior can distort interventional predictions and yield an ill-conditioned covariance requiring regularization (e.g., diagonal jitter). Therefore, the causal prior should *guide*, not *dominate*. A dominating prior would have hurt RESCUE under misspecification. Instead, VoCI-HVKG adjusts reliance on the prior through causal trust $\widehat{\omega}$, while flexible GP hyperparameters allow erroneous causal information to be corrected or overridden. In unconstrained problems, a misleading low-fidelity evaluation hurts objective estimation, and the optimizer can recover by sampling higher fidelities. Under constraints, it must learn both the Pareto front and feasible region, and distorted low-fidelity feasibility estimates can waste budget on configurations infeasible at the target fidelity. RESCUE reduces this failure mode since the causal prior also informs the constraints in the MF-CGP.

We use DoWhy (Blöbaum et al., 2024) for causal inference and train a neural-network proxy on CPM interventional samples for PyTorch (Paszke et al., 2019) integration and GPU acceleration. Directly calling DoWhy would avoid this proxy, but was empirically too slow in our setting. The proxy accurately approximates *do*-calculus estimates (avg. $R^2 = 0.97$) with a 2.4×10^4 speedup (Appendix A.6). Any bias introduced by this approximation is included in ξ of Assumption 5.2.

RESCUE depends on the availability and quality of causal information. The causal graph may come from domain knowledge, causal discovery, or a combination of both. The NP-hard complexity of causal discovery (Chickering et al., 2004) implies that the identified causal model may not always represent the ground-truth causal relationships among variables. It is therefore important to recognize potential discrepancies between the discovered causal structure and the actual one.

Future work could reduce the search space using causal-effect estimates to prune configurations with negligible effects on target-fidelity objectives or constraints, and replace the GP with scalable probabilistic models for high-dimensional, conditional search spaces.

IMPACT STATEMENT

RESCUE supports decisions on real systems from a causal prior learned from historical data. A misspecified causal graph, an inaccurate proxy or a wrong constraint model can recommend unsafe configurations, so the causal trust and the constraint estimates should be monitored before deployment.

AI USE STATEMENT

In this work, we used generative AI tools to proofread the paper for grammar, document the code, write the scripts that automate the experiments and create the figures. We also used them to check the mathematical proofs, to provide feedback on the experiment design, and to assist with implementing the method. We have not used generative AI tools to develop the theoretical framework, formulate mathematical claims, propose hypotheses, generate synthetic data sets or clean data sets, and translation and qualitative data analysis are not applicable to this work. We have reviewed all AI-assisted work. The authors verified every proof step, tested all AI-assisted code and validated the results against the logged runs. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

We provide the code at `https://anonymous.4open.science/r/rescue-527D` together with the observational datasets, the initial designs and the run histories of all runs and the scripts that reproduce the results. We log every run in Weights & Biases, including the per-iteration metrics and a checkpoint of each optimization step with the training data, the surrogate, the discovered DAG, the CPM and the recommended Pareto set. Appendix A provides the initial sampling settings and the optimization budget of each problem (Table 2). It also provides the implementation details of all methods and the definition of each problem with its cost function. Appendix A.7 describes the multi-fidelity robot navigation benchmark. The ablation and proxy quality experiments are specified in Appendices A.5 and A.6. The assumptions of the theoretical results are stated in Section 5 and the complete proofs are in Appendix B.

REFERENCES

Virginia Aglietti, Xiaoyu Lu, Andrei Paleyes, and Javier González. Causal Bayesian optimization. In *International Conference on Artificial Intelligence and Statistics*, pp. 3155–3164. PMLR, 2020.

Virginia Aglietti, Alan Malek, Ira Ktena, and Silvia Chiappa. Constrained causal Bayesian optimization. In *International Conference on Machine Learning*, pp. 304–321. PMLR, 2023.

Maximilian Balandat, Brian Karrer, Daniel Jiang, Samuel Daulton, Ben Letham, Andrew G Wilson, and Eytan Bakshy. Botorch: A framework for efficient Monte-Carlo Bayesian optimization. *Advances in neural information processing systems*, 33:21524–21538, 2020.

Syrine Belakaria, Aryan Deshwal, and Janardhan Rao Doppa. Multi-fidelity multi-objective Bayesian optimization: An output space entropy search approach. In *Proceedings of the AAAI Conference on artificial intelligence*, volume 34, pp. 10035–10043, 2020.

Shriya Bhatija, Paul-David Zuercher, Jakob Thumm, and Thomas Bohné. Multi-objective causal bayesian optimization. In *Proceedings of the 42nd International Conference on Machine Learning*, 2025.

Patrick Blöbaum, Peter Götz, Kailash Budhathoki, Atalanti A. Mastakouri, and Dominik Janzing. DoWhy-GCM: An extension of DoWhy for causal inference in graphical causal models. *Journal of Machine Learning Research*, 25(147):1–7, 2024. URL `http://jmlr.org/papers/v25/22-1258.html`.

Ilija Bogunovic and Andreas Krause. Misspecified gaussian process bandit optimization. *Advances in neural information processing systems*, 34:3004–3015, 2021.

Max Chickering, David Heckerman, and Chris Meek. Large-sample learning of Bayesian networks is NP-hard. *Journal of Machine Learning Research*, 5:1287–1330, 2004.

Sayak Ray Chowdhury and Aditya Gopalan. On kernelized multi-armed bandits. In *International Conference on Machine Learning*, pp. 844–853. PMLR, 2017.

Sam Daulton, Maximilian Balandat, and Eytan Bakshy. Hypervolume knowledge gradient: a lookahead approach for multi-objective Bayesian optimization with partial information. In *International Conference on Machine Learning*, pp. 7167–7204. PMLR, 2023.

Samuel Daulton, Maximilian Balandat, and Eytan Bakshy. Differentiable expected hypervolume improvement for parallel multi-objective Bayesian optimization. *Advances in neural information processing systems*, 33:9851–9864, 2020.

Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and TAMT Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197, 2002.

David Galles and Judea Pearl. Testing identifiability of causal effects. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, UAI'95, pp. 185–195, 1995.

Youwei He, Jinju Sun, Peng Song, and Xuesong Wang. Variable-fidelity hypervolume-based expected improvement criteria for multi-objective efficient global optimization of expensive functions. *Engineering with Computers*, 38(4):3663–3689, 2022.

Daniel Hernández-Lobato, Jose Hernandez-Lobato, Amar Shah, and Ryan Adams. Predictive entropy search for multi-objective Bayesian optimization. In *International conference on machine learning*, pp. 1492–1501. PMLR, 2016.

Md Abir Hossen, Sonam Kharade, Bradley Schmerl, Javier Cámara, Jason M. O'Kane, Ellen C. Czaplinski, Katherine A. Dzurilla, David Garlan, and Pooyan Jamshidi. CaRE: Finding root causes of configuration issues in highly-configurable robots. *IEEE Robotics and Automation Letters*, 8(7): 4115–4122, 2023. doi: 10.1109/LRA.2023.3280810.

Md Abir Hossen, Sonam Kharade, Jason M. O'Kane, Bradley Schmerl, David Garlan, and Pooyan Jamshidi. CURE: Simulation-augmented auto-tuning in robotics. *IEEE Transactions on Robotics*, pp. 1–17, 2025. doi: 10.1109/TRO.2025.3548546.

Carl Hvarfner, Erik Orm Hellsten, and Luigi Nardi. Vanilla Bayesian optimization performs great in high dimensions. In *International Conference on Machine Learning*. PMLR, 2024.

Md Shahriar Iqbal, Rahul Krishna, Mohammad Ali Javidian, Baishakhi Ray, and Pooyan Jamshidi. Unicorn: Reasoning about configurable system performance through the lens of causality. In *Proceedings of the Seventeenth European Conference on Computer Systems (EuroSys)*, 2022.

Md Shahriar Iqbal, Ziyuan Zhong, Iftakhar Ahmad, Baishakhi Ray, and Pooyan Jamshidi. CAMEO: A causal transfer learning approach for performance optimization of configurable computer systems. In *Proceedings of the 2023 ACM Symposium on Cloud Computing*, pp. 555–571, 2023.

Faran Irshad, Stefan Karsch, and Andreas Döpp. Leveraging trust for joint multi-objective and multi-fidelity optimization. *Machine Learning: Science and Technology*, 5(1):015056, 2024.

Hisao Ishibuchi, Hiroyuki Masuda, Yuki Tanigaki, and Yusuke Nojima. Modified distance calculation in generational distance and inverted generational distance. In *International Conference on Evolutionary Multi-Criterion Optimization*, pp. 110–125. Springer, 2015.

Hisao Ishibuchi, Ryo Imada, Yu Setoguchi, and Yusuke Nojima. How to specify a reference point in hypervolume calculation for fair performance comparison. *Evolutionary Computation*, 26(3): 411–440, 2018.

Kirthevasan Kandasamy, Gautam Dasarathy, Junier B Oliva, Jeff Schneider, and Barnabas Poczos. Gaussian process bandit optimisation with multi-fidelity evaluations. In *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016.

Rémi Lam, Douglas L Allaire, and Karen E Willcox. Multifidelity optimization using statistical surrogate modeling for non-hierarchical information sources. In *56th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, pp. 0143, 2015.

Shibo Li, Wei Xing, Robert Kirby, and Shandian Zhe. Multi-fidelity Bayesian optimization via deep neural networks. *Advances in Neural Information Processing Systems*, 33:8521–8531, 2020.

Steven Macenski, Francisco Martin, Ruffin White, and Jonatan Ginés Clavero. The marathon 2: A navigation system. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020.

Steven Macenski, Tully Foote, Brian Gerkey, Chris Lalancette, and William Woodall. Robot operating system 2: Design, architecture, and uses in the wild. *Science Robotics*, 7(66):eabm6074, 2022. doi: 10.1126/scirobotics.abm6074. URL https://www.science.org/doi/abs/10.1126/scirobotics.abm6074.

Petrus Mikkola, Julien Martinelli, Louis Filstroff, and Samuel Kaski. Multi-fidelity Bayesian optimization with unreliable information sources. In *International Conference on Artificial Intelligence and Statistics*, pp. 7425–7454. PMLR, 2023.

Henry B Moss, David S Leslie, Javier Gonzalez, and Paul Rayson. GIBBON: General-purpose information-based Bayesian optimisation. *Journal of Machine Learning Research*, 22(235):1–49, 2021.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.

Judea Pearl. *Causality*. Cambridge university press, 2009.

Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of causal inference: foundations and learning algorithms*. The MIT Press, 2017.

Florian Pfisterer, Lennart Schneider, Julia Moosbauer, Martin Binder, and Bernd Bischl. YAHPO Gym–an efficient multi-objective multi-fidelity benchmark for hyperparameter optimization. In *International Conference on Automated Machine Learning*, pp. 3–1. PMLR, 2022.

Matthias Poloczek, Jialei Wang, and Peter Frazier. Multi-information source optimization. *Advances in neural information processing systems*, 30, 2017.

Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. Catboost: unbiased boosting with categorical features. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL https://proceedings.neurips.cc/paper_files/paper/2018/file/14491b756b3a51daac41c24863285549-Paper.pdf.

Shohei Shimizu, Takanori Inazumi, Yasuhiro Sogawa, Aapo Hyvarinen, Yoshinobu Kawahara, Takashi Washio, Patrik O Hoyer, Kenneth Bollen, and Patrik Hoyer. DirectLiNGAM: A direct method for learning a linear non-Gaussian structural equation model. *Journal of Machine Learning Research-JMLR*, 12(Apr):1225–1248, 2011.

Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, prediction, and search*. MIT press, 2001.

Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995*, 2009.

Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *International Conference on Machine Learning*, pp. 1015–1022. PMLR, 2010.

S. Sundararajan and S. Sathiya Keerthi. Predictive approaches for choosing hyperparameters in Gaussian processes. *Neural Computation*, 13(5):1103–1118, 2001.

Shion Takeno, Hitoshi Fukuoka, Yuhki Tsukada, Toshiyuki Koyama, Motoki Shiga, Ichiro Takeuchi, and Masayuki Karasuyama. Multi-fidelity Bayesian optimization with max-value entropy search and its parallelization. In *International Conference on Machine Learning*, pp. 9334–9345. PMLR, 2020.

Petru Tighineanu, Kathrin Skubch, Paul Baireuther, Attila Reiss, Felix Berkenkamp, and Julia Vinogradska. Transfer learning with Gaussian processes for Bayesian optimization. In *International Conference on Artificial Intelligence and Statistics*, pp. 6152–6181. PMLR, 2022.

Dongxia Wu, Ruijia Niu, Matteo Chinazzi, Yian Ma, and Rose Yu. Disentangled multi-fidelity deep Bayesian active learning. In *International Conference on Machine Learning*, pp. 37624–37634. PMLR, 2023.

Yuling Yao, Aki Vehtari, Daniel Simpson, and Andrew Gelman. Using stacking to average Bayesian predictive distributions (with discussion). *Bayesian Analysis*, 13(3):917–1007, 2018.

Jakob Zeitler. Towards MFACBO: Multi-fidelity abstraction causal bayesian optimization. In *UAI Workshop on Causal Abstractions and Representations*, 2025.

Qi Zhou, Min Zhao, Jiexiang Hu, and Mengying Ma. *Multi-fidelity Surrogates Modeling, Optimization and Applications.* Springer Singapore, 2023. URL https://doi.org/10.1007/978-981-19-7210-2.

APPENDIX CONTENTS

A Experiment Details **15**

A.1 Cost-Aware Initial Sampling 15

A.2 Implementation of Acquisition Functions and Models 15

A.3 Problem Details 16

A.4 Additional Results 18

A.5 Ablation Experiments 20

A.6 Quality of causal model proxy 21

A.7 Multi-fidelity mobile robot navigation benchmark 22

B Proofs **30**

B.1 Proof of Lemma 5.6 30

B.2 Proof of Lemma 5.7 30

B.3 A feasibility sandwich 31

B.4 Proof of Theorem 5.8 32

B.5 Proof of Corollary 5.9 33

B.6 A sublinear rate under periodic target-fidelity exploration 33

B.7 Discussion 36

NOTATIONS

Table 1: Mathematical notations.

Description	Notation	Description	Notation
Configuration space	$\mathbb{X}$	Pareto frontier at target fidelity	$\mathbb{P}$
Feasible configuration space	$\bar{\mathbb{X}} \subseteq \mathbb{X}$	Recommended Pareto set	$\widehat{\mathbb{P}}_T$
Estimated feasible region	$\widehat{\mathbb{X}}_T \subseteq \mathbb{X}$	Directed acyclic graph (DAG)	$\mathcal{G}$
Fidelity space	$\mathbb{S}$	Causal performance model (CPM)	$\mathcal{C}$
Configuration	$\boldsymbol{x} \in \mathbb{X}$	Causal prior	$\bar{\boldsymbol{f}}_{\mathrm{do}}(\boldsymbol{x}, s)$
Configuration-fidelity pair	$\boldsymbol{\chi} := (\boldsymbol{x}, s)$	Observational dataset	$\widehat{\mathcal{D}}$
Fidelity	$s \in \mathbb{S}$	Interventional dataset	$\mathcal{D}$
Target fidelity	$s_\diamond \in \mathbb{S}$	Per-output reliability	$\boldsymbol{\omega} \in [0,1]^J$
Number of configuration options	d	CPM error	ξ
Number of objectives	M	Maximum information gain	γ_t
Number of constraints	Q	Confidence parameter	β_t
Total number of outputs	$J = M + Q$	Target-fidelity posterior uncertainty	$\bar{\sigma}_T$
Objective value	$\boldsymbol{y}$	Exploration cycle	N_{e}
Constraint functions	$h_q(\boldsymbol{x}, s) \geq 0$	Hypervolume-stability constant	L
Cross-fidelity bias	$\boldsymbol{\delta}(\boldsymbol{x}, s)$	Constraint-margin constant	L'
Total budget	Λ	Reference point	$\boldsymbol{r}$
Total number of interventions	T	RKHS residual of the causal prior	$\boldsymbol{w}$
		Candidate set of the value function	$\mathbb{B}$

Algorithm 2 Cost-Aware Initial Sampling

Require: $\mathbb{X}$, $\mathbb{S}$, $c(\cdot,\cdot)$, sampling budget Λ_{init} in target-fidelity evaluations, maximum design size $N_{\max}$, highest fidelity $s_{\max}$
Ensure: Initial dataset $\mathcal{D}_0 = \{(\boldsymbol{x}_i, s_i, \boldsymbol{y}_i)\}_{i=1}^{N}$
1: Initialize $\mathcal{D}_0 = \emptyset$, $C_0 = 0$, $i = 0$
2: **while** $C_i < \Lambda_{\text{init}}\,\bar{c}_{s_\diamond}$ and $i < N_{\max}$ **do**
3: Sample configuration $\mathbf{x} \sim \text{Unif}(\mathbb{X})$
4: Sample fidelity $s = \text{CDF}^{-1}(u)$ using Eq. (6) restricted to $[0, s_{\max}]$, where u $\sim \text{Unif}(0,1)$
5: **if** $C_i + c(\boldsymbol{x}, s) > \Lambda_{\text{init}}\,\bar{c}_{s_\diamond}$ **then**
6: **break**
7: **end if**
8: $i \leftarrow i+1$, $(\boldsymbol{x}_i, s_i) \leftarrow (\boldsymbol{x}, s)$
9: Evaluate $\boldsymbol{y}_i = \bar{\boldsymbol{f}}(\boldsymbol{x}_i, s_i)$
10: $\mathcal{D}_{0:i} \leftarrow \mathcal{D}_{0:i-1} \cup \{(\boldsymbol{x}_i, s_i, \boldsymbol{y}_i)\}$
11: $C_i \leftarrow C_{i-1} + c(\boldsymbol{x}_i, s_i)$
12: **end while**
13: **return** $\mathcal{D}_0$

A Experiment Details

A.1 Cost-Aware Initial Sampling

To initialize $\mathcal{D}_0 = \{(\boldsymbol{x}_i, s_i, \boldsymbol{y}_i)\}_{i=1}^{N}$ (line 1 of Algorithm 1), we sample configurations and fidelities using Algorithm 2, where the CDF is:

$$\text{CDF}(s) = \begin{cases} \dfrac{\int_0^s c(\boldsymbol{x}, \tilde{s})^{-1} d\tilde{s}}{\int_0^1 c(\boldsymbol{x}, \tilde{s})^{-1} d\tilde{s}} & \text{(continuous fidelities)} \\ \dfrac{\sum_{j:s_j \le s} c(\boldsymbol{x}, s_j)^{-1}}{\sum_{j=1}^{|\mathbb{S}|} c(\boldsymbol{x}, s_j)^{-1}} & \text{(discrete fidelities),} \end{cases} \tag{6}$$

and $p(s) \propto c(\boldsymbol{x}, s)^{-1}$, so that lower-cost fidelities are sampled more frequently, increasing sample diversity under the initialization budget. The initial design is capped at $N_{\max}$ points and at a highest fidelity $s_{\max}$, where $N_{\max} = 2(d+1)$, except for robot navigation, where $N_{\max} = 0.5(d+1)$ to limit the number of real-robot runs. The sampling budget Λ_{init} is given in target-fidelity evaluations, and Algorithm 2 charges it in cost units through the median cost of a target-fidelity evaluation, $\bar{c}_{s_\diamond}$. Within each run, all methods start from the same initial trials. Table 2 lists the resulting initial trials of each problem. The cost of the initial trials is not counted against Λ. Table 3 lists the data of each problem.

A.2 Implementation of Acquisition Functions and Models

We use the BoTorch implementations[6] of qEHVI (Daulton et al., 2020), HVKG (Daulton et al., 2023), and MOMF (Irshad et al., 2024). qEHVI is a single-fidelity acquisition function. Evaluated at $s_\diamond$ only, each of its interventions on Branin-Currin and Park, for example, costs $\exp(4.8) \approx 121$ under $c(\boldsymbol{x}, s) = \exp(4.8s)$, whereas an intervention at the lowest fidelity costs 1, so its regret over cost is determined by this ratio from the first intervention at any budget. We therefore add s to the design variables of qEHVI, which lets it select the fidelity of each intervention. The baselines use the multi-task GP of section 4.1 without the causal prior mean and the causal covariance term,

$$\bar{\boldsymbol{f}}(\boldsymbol{\chi}) \sim \text{GP}\big(\boldsymbol{c},\, k_{j,j'}(\boldsymbol{\chi}, \boldsymbol{\chi}')\big), \qquad k_{j,j'}(\boldsymbol{\chi}, \boldsymbol{\chi}') = k(\boldsymbol{\chi}, \boldsymbol{\chi}') \otimes k_{\text{obj}}(j, j'),$$

where $\boldsymbol{c}$ is a learned constant mean, k is an ARD RBF kernel, and k_{obj} is the index kernel of section 4.1. All methods use the dimension-scaled log-normal lengthscale prior of Hvarfner et al. (2024), standardized outputs, and a learned noise level, and their hyperparameters are fit by maximizing the exact marginal likelihood. For all MC acquisition functions, we use quasi-random base samples

[6] https://github.com/meta-pytorch/botorch

Table 2: Initial trials of each problem. N_0 is the number of initial trials and their cost is the range over the 20 seeds in the units of Λ.

Problem	$s_{\max}$	N_0	Initial cost
Branin-Currin	0.35	6	6.5–13.1
Park	0.35	10	11.9–20.8
Synthetic SCM	0.2	10	4.7–6.9
Robot navigation	0.20	13	1440–2152
Constrained XGBoost HPO	1.0	24–26	66–120

and sample average approximation (Balandat et al., 2020). All acquisition functions are optimized with L-BFGS-B; HVKG and VoCI-HVKG in their one-shot form from a single starting point, as in Daulton et al. (2023). To determine the Pareto frontier of the posterior mean, we use NSGA-II (Deb et al., 2002) with a population of 250 for 100 generations. Table 4 lists the acquisition settings.

For RESCUE, the DAG is learned from the observational data with the PC algorithm (Spirtes et al., 2001) for Branin-Currin, Park ($\alpha = 0.52$ from a schedule that decreases with $|\widehat{\mathcal{D}}|$) and robot navigation ($\alpha = 0.05$), and with DirectLiNGAM (Shimizu et al., 2011) for Synthetic SCM, and constrained XGBoost HPO with structural constraints that output variables (objective and constraint) do not cause a configuration option. PC orients edges with the structural constraints as background knowledge, and edges it leaves undirected are dropped. In our experiments, $\widehat{\mathcal{D}}$ is sampled uniformly at random from each problem as a stand-in for historical data. In practice, $\widehat{\mathcal{D}}$ comes from system logs, such as robot telemetry or the trial history of an AutoML service, and the causal discovery step remains unchanged. The structural equations are fit with DoWhy (Blöbaum et al., 2024), and the causal prior is computed by *do*-calculus.

The causal prior is expensive to evaluate by *do*-calculus inside the acquisition function. We distill it into a proxy network. The interventional means and standard deviations are computed at $N_{\mathrm{do}} = 500$ configuration–fidelity pairs (200 for robot navigation). For each output, we fit two GPs, one on the means and one on the standard deviations. A two-layer MLP with 64 hidden units per layer, ReLU activations, a mean head, and a softplus standard-deviation head is trained on 16,384 Sobol points labeled by these GPs. It minimizes the mean squared error of the standardized mean and standard deviation with Adam (learning rate 10^{-3}, at most 8000 epochs, early stopping on the training loss). The proxy provides $\bar{\boldsymbol{f}}_{\mathrm{do}}$ and $\widehat{\boldsymbol{\sigma}}$ to the MF-CGP in the space of the standardized training targets and is frozen while the MF-CGP hyperparameters are fit. We report its accuracy in Appendix A.6.

A.3 Problem Details

Table 5 lists the reference points of all problems.

A.3.1 Synthetic problems

Branin-Currin and Park We use the modified multi-fidelity, multi-objective Branin-Currin and Park function from Irshad et al. (2024). The configuration space is given in Table 10a and Table 10b, and the discovered DAG is shown in Fig. 14.

Synthetic SCM We use the structural equations from Bhatija et al. (2025) (MO-CBO problem, Synthetic-2)[7]. Since there is no multi-fidelity variant of this problem, we modify the mechanisms to incorporate an additional fidelity input dimension following Zeitler (2025):

$$
\begin{aligned}
&X_4 = U + U_{X_4}^2, \quad X_1 = X_4/2 + U_{X_1}^2, \quad X_2 = U_{X_2}^2, \quad X_3 = U_{X_3}^2, \\
&Y_1 = \log(1 + X_1^2) + \big(2 - 0.4(1-s)\big)X_2^2 - X_1 X_2 U/2 + U_{Y_1}^3, \\
&Y_2 = \sin\big((0.7 + 0.3s)X_2^2\big) - X_3^2 - \big(1 - 0.3(1-s)\big)X_2 X_3 + 50 + U_{Y_2}^3,
\end{aligned}
$$

where $U \in \{-4, 4\}$ with $P(U = -4) = P(U = 4) = 0.5$ is an unobserved confounder of X_4 and Y_1, and $U_{X_i}, U_{Y_i} \sim \mathcal{N}(0, 0.5)$. The main difference is the fidelity parameter s in the

[7] https://github.com/ShriyaBhatija/MO-CBO

Table 3: Data of each problem. $\widehat{\mathcal{D}}$ is used for causal discovery only. The initial trials are shared by all methods within a seed. The held-out evaluation data serve to compute the metrics, and no method has access to them.

Problem	**Historical $\widehat{\mathcal{D}}$**	**Benchmark-training data**	**BO data**	**Held-out evaluation data**
Branin-Currin	500 samples per seed	None, analytic functions	6 initial trials and the queries within $\Lambda = 20$	True Pareto front of the analytic functions at $s_\diamond$
Park	500 samples per seed	None, analytic functions	10 initial trials and the queries within $\Lambda = 10$	True Pareto front at $s_\diamond$
Synthetic SCM	500 samples per seed	None, structural equations	10 initial trials and the queries within $\Lambda = 20$	NSGA-II front of the structural equations at $s_\diamond$
Robot navigation	1,000 samples per seed from the benchmark model	240 real-robot runs, 80% of 300	13 initial trials and the queries within $\Lambda = 2500$	60 real-robot runs for the benchmark model, NSGA-II front of the model at $s_\diamond$
Constrained XGBoost HPO	1,000 samples per seed from YAHPO Gym	YAHPO Gym surrogate, trained by its authors	24–26 initial trials and the queries within $\Lambda = 20$	NSGA-II front of YAHPO Gym at $s_\diamond$

Table 4: Acquisition function and optimization settings.

Method	**Parameter**	**Description**	**Value**
All	`raw_samples`	Initial candidates of the acquisition optimization	512
	`maxiter`	L-BFGS-B iterations	200
	`eta`	Feasibility sigmoid temperature (constrained problems)	10^{-3}
qEHVI, MOMF	`mc_samples`	MC samples	128
	`num_restarts`	L-BFGS-B starting points	10
HVKG, VoCI-HVKG	`num_fantasies`	Fantasy samples N_f	8 (RN: 6)
	`num_pareto`	Candidate Pareto set size $n_{\mathbb{P}}$	10 (RN: 8)
	`num_restarts`	L-BFGS-B starting points	1
	`use_posterior_mean`	ν evaluated on the posterior mean	True
VoCI-HVKG	`fit_restarts`	Marginal likelihood restarts of $\mathcal{M}_\mathbf{0}$ and $\mathcal{M}_\mathbf{1}$	4
	`prune_threshold`	Causal trust below which a component is discarded	0.02
	`update_weights`	$\widehat{\boldsymbol{\omega}}$ updated in the lookahead	True

coefficients, which recovers the original mechanisms when $s = 1$. X_4 affects Y_1 only through X_1, so it has no effect once X_1 is intervened on. The nonlinear mechanisms violate the assumptions of DirectLiNGAM, so this problem evaluates RESCUE under a misspecified causal discovery step. We minimize Y_1 and Y_2. Each evaluation intervenes on all four configuration options at unit cost per option, as in Bhatija et al. (2025), scaled by the fidelity, $c(\boldsymbol{x}, s) = 4\,(0.1 + 0.9s)$. The configuration space is given in Table 10c and the discovered DAG is shown in Fig. 14. The six edges of the structural equations into Y_1 and Y_2 are recovered in every run, and no spurious edge appears in more than 2 of the 20 runs.

Table 5: Reference points for all benchmark problems (assuming minimization of all objectives). Branin-Currin and Park use the reference point of Irshad et al. (2024). For the other problems, the true Pareto front at the target fidelity is the non-dominated union of 20 NSGA-II runs on the true objectives with a population of $1{,}000$ for 100 generations. The reference point is the nadir of this front offset by 10% of its range in every objective (Ishibuchi et al., 2018).

Problem	**Reference point (r)**
Branin-Currin	$(0, 0)$
Park	$(0, 0)$
Synthetic SCM	$(3.5374, 2.3558)$
Robot navigation	$(0.0255, 4.8319)$
Constrained XGBoost HPO	$(0.6824, 2.3928)$

A.3.2 Real-world problems

Robot navigation Following the experimental setup of Hossen et al. (2025), we deploy mobile robots in a controlled indoor environment. The robots are operated using ROS 2 (Macenski et al., 2022) and the Nav2 (Macenski et al., 2020) navigation stack for localization, path planning, and trajectory tracking. We consider point-to-point navigation tasks and record navigation performance across fidelity levels. Specifically, we measure (i) $\text{Energy per meter} = \frac{\text{Energy}_{\text{total}}}{\text{Distance traveled}}$, (ii) Task execution time, (iii) $\text{Task completion rate} = \frac{\sum \text{Tasks}_{\text{completed}}}{\sum \text{Tasks}}$, and (iv) a physics-based Collision-risk score defined as follows:

$$\text{Collision-risk score} = 10(0.2{\cdot}r_{\text{speed}}{+}0.3{\cdot}r_{\text{safety}}{+}0.2{\cdot}r_{\text{reaction}}{+}0.1{\cdot}r_{\text{perception}}{+}0.15{\cdot}r_{\text{goal}}{+}0.05{\cdot}r_{\text{sampling}}),$$

where

$$r_{\text{speed}} = \text{clip}\left(\frac{\text{max_vel_x}}{2\,(|\,\text{decel_lim_x}\,| + 0.01)}, 0, 1\right), \quad r_{\text{safety}} = \frac{1}{1 + \frac{\text{local_inflation_radius}}{\text{max_vel_x} + 0.01}},$$

$$r_{\text{reaction}} = \frac{1}{1 + \frac{\text{local_inflation_radius}}{\text{max_vel_x} \cdot \text{sim_time} + 0.01}}, \quad r_{\text{perception}} = \text{clip}\left(\frac{\text{local_resolution}}{0.5\,(\text{local_inflation_radius} + 0.01)}, 0, 1\right),$$

$$r_{\text{goal}} = \text{clip}\left(\frac{\text{GoalAlign. scale} + \text{GoalDist. scale}}{3\,(\text{BaseObstacle. scale} + 0.01)}, 0, 1\right), \quad r_{\text{sampling}} = \frac{1}{1 + \frac{\text{vx_samples} \cdot \text{vtheta_samples}}{200}}.$$

Note that $\text{clip}(x, 0, 1) := \max(0, \min(x, 1))$. We minimize energy per meter and the collision-risk score subject to a task completion rate of at least 0.95. The MF variant of the collision-risk score is

$$\text{Collision-risk score}(s) = 10\Big(1 - \big(1 - \tfrac{\text{Collision-risk score}}{10}\big)^{\text{Distance traveled}(s)/\text{Distance traveled}(s_\diamond)}\Big).$$

The evaluation cost is the sum of the task completion time and a setup time of 70 seconds. Since the cost is known only after the run, a GP on the log cost fit to evaluations of the benchmark model predicts it for the acquisition, and the budget counts the measured cost. The configuration space is given in Table 10e and the discovered DAG is shown in Fig. 15.

Constrained XGBoost HPO We use the YAHPO Gym (Pfisterer et al., 2022) multi-fidelity, multi-objective HPO benchmark (Scenario: iaml_xgboost, Instance: 41146). The objectives are the logloss and the main effect complexity of the model (mec), both to be minimized while enforcing that the memory usage during training (ramtrain) is at most 15.64 MB. The fidelity is the fraction of the training set and the evaluation cost is the training time (timetrain), modeled as a function of the fidelity fit to the median timetrain of the benchmark. The configuration space is given in Table 10d and the discovered DAG is shown in Fig. 16.

A.4 Additional Results

Fig. 6 reports attainment probability as a function of normalized distance to the true Pareto front over all iterations. Across all problems, RESCUE achieves higher attainment probabilities at smaller distances than the baselines. The empirical attainment functions in Fig. 12 complement these results,

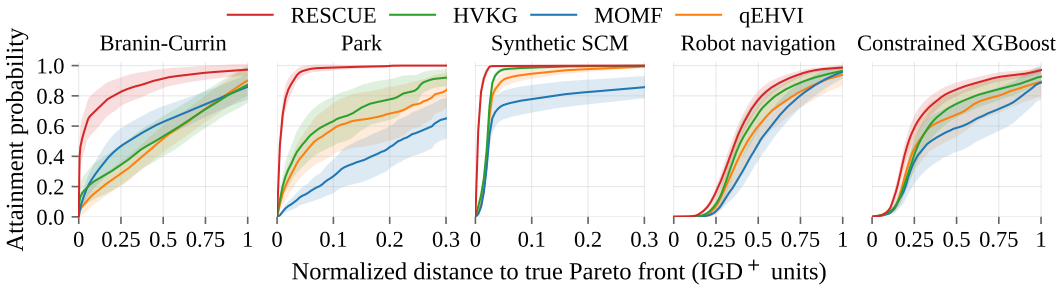


Figure 6: Attainment probability of the recommended Pareto set against the normalized distance to the true Pareto front.

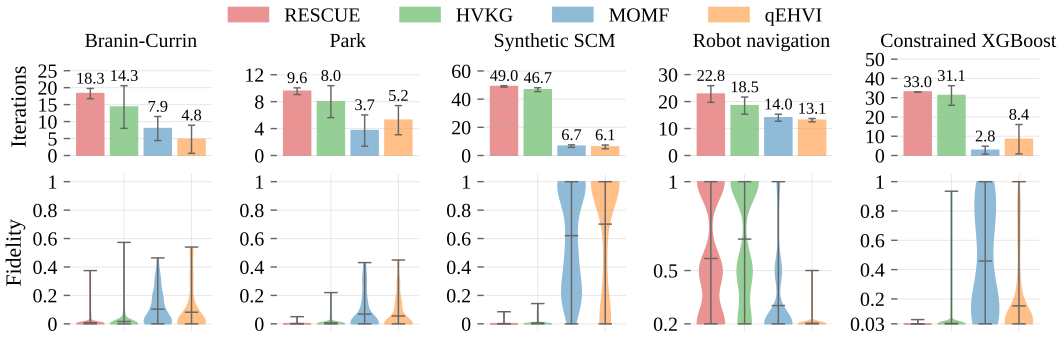


Figure 7: Fidelity queries and iteration counts under a fixed budget. For each problem, the top panel reports the number of iterations achieved by each method, and the bottom panel shows the empirical distribution of queried fidelities. Each violin shows the fidelity distribution with a vertical min–max range line and a mean marker.

showing attainment regions that extend closer to the true front in objective space. HVKG is the strongest baseline, while qEHVI, despite its single-fidelity acquisition function, outperforms MOMF on all problems except Branin-Currin.

Why does RESCUE work better? To understand why RESCUE achieves superior optimization performance compared to the baselines, we examine how it uses the budget. Fig. 7 shows that RESCUE uses cheaper fidelities more often than the baselines. HVKG performs on par with RESCUE, as both methods implement similar acquisition functions, but it never surpasses RESCUE. In contrast, qEHVI and MOMF consume the budget much faster with high fidelity queries. This behavior is expected for qEHVI, whose acquisition function has no fidelity model, and explains why MOMF performs similarly or worse than qEHVI despite being an MF method. Robot navigation differs since its cost depends on the configuration as well as the fidelity. Here qEHVI and MOMF query the lowest fidelity almost exclusively but select configurations with long execution times, while RESCUE and HVKG select configurations with lower execution times at higher fidelities and achieve more iterations.

Runtime All experiments were run on a single machine with an AMD Ryzen Threadripper 3960X CPU, an NVIDIA RTX A6000 GPU, 128 GB memory, and Ubuntu 20.04 with Python 3.11, BoTorch 0.18.1 and DoWhy 0.13. Fig. 8 reports the runtime per iteration of each method, excluding the time to obtain measurements. Across the five benchmarks, qEHVI and MOMF have the smallest runtimes, while RESCUE and HVKG are slower. RESCUE remains in the same order of magnitude as HVKG. The higher runtime of RESCUE and HVKG arises from computing the (VoCI-)HVKG acquisition with Monte Carlo fantasies. In RESCUE this cost is further increased by fitting the reliability mixture (Eq. (2)) at each iteration and by learning the CPM once per run.

Experiment statistics Table 6 reports the area-under-regret (AUR) curve. For each problem and baseline, the AUR is obtained by integrating the instantaneous regret over the total budget using the

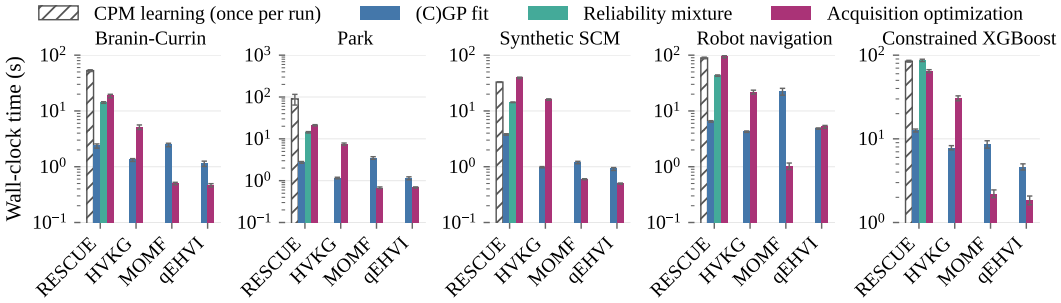


Figure 8: Optimization runtime for each problem and method across components. For every problem–method pair, we repeated the experiment with 20 independent seeds. Bars represent the mean wall-clock seconds per iteration with standard errors. The CPM of RESCUE is learned once per run (Algorithm 1).

Table 6: Area-under-regret (AUR) curve comparison between RESCUE and baselines. Blue cells in the p-value column denote $p < 0.05$, and green/red cells in the Gain column indicate improvements / degradations of RESCUE over the corresponding baseline.

Problem	Baseline	AUR ↓	p-value	t-stat.	Cohen's d	RESCUE	
						Gain %	AUR ↓
Branin-Currin ($d = 2, M = 2$)	qEHVI	-12.39 ± 4.62	9.05E−7	−7.12	−1.59	98.07	-24.55 ± 7.49
	MOMF	-14.67 ± 4.84	7.79E−5	−5.01	−1.12	67.31	
	HVKG	-14.03 ± 6.78	3.12E−4	−4.39	−0.98	74.97	
Park ($d = 4, M = 2$)	qEHVI	-13.63 ± 3.21	7.36E−9	−9.79	−2.19	62.21	-22.11 ± 2.61
	MOMF	-10.39 ± 1.09	9.50E−14	−18.83	−4.21	112.77	
	HVKG	-14.69 ± 3.79	2.96E−7	−7.70	−1.72	50.51	
Synthetic SCM ($d = 4, M = 2$)	qEHVI	-7.58 ± 3.56	1.57E−11	−14.13	−3.16	185.78	-21.66 ± 2.97
	MOMF	-8.37 ± 2.35	8.04E−12	−14.68	−3.28	158.66	
	HVKG	-10.75 ± 2.30	9.03E−11	−12.77	−2.85	101.56	
Robot navigation ($d = 25, M = 2, Q = 1$)	qEHVI	-4025.31 ± 184.68	1.55E−4	−4.70	−1.05	5.28	-4237.67 ± 131.03
	MOMF	-3961.45 ± 134.62	1.62E−7	−8.02	−1.79	6.97	
	HVKG	-4019.72 ± 123.52	1.95E−5	−5.64	−1.26	5.42	
Constrained XGBoost HPO ($d = 12, M = 2, Q = 1$)	qEHVI	-5.01 ± 2.42	2.73E−1	−1.13	−0.25	16.29	-5.83 ± 1.41
	MOMF	-4.07 ± 2.79	2.41E−2	−2.45	−0.55	43.27	
	HVKG	-5.07 ± 1.78	2.12E−1	−1.29	−0.29	15.13	

trapezoidal rule. For each baseline, we report the paired t-test p-value, t-statistic, Cohen's d, and the corresponding relative AUR gain of RESCUE. Across the five benchmarks, RESCUE reduces AUR against every baseline, with statistically significant gains in 13 of the 15 comparisons, the exceptions being qEHVI and HVKG on constrained XGBoost HPO. Overall, RESCUE achieves an improvement of **74%** over qEHVI, **78%** over MOMF, and **50%** over HVKG.

A.5 Ablation Experiments

Table 7 describes the ablation experiment design for RESCUE. Every variant shares the initial data, budget, and evaluation protocol of the main experiments. Under the misspecified causal prior, the CPM is learned from $\widehat{\mathcal{D}}$ in which each outcome is paired with a different configuration–fidelity pair, so the CPM learns an incorrect dependence of the outcomes on configurations and fidelities and ξ is large. By Theorem 5.8, an auxiliary source affects the regret bound only through ξ, so this ablation also tests an unreliable auxiliary source. Fig. 9 reports the best log HV regret of the ablation variants and Fig. 13 shows their empirical attainment functions of the model selected Pareto fronts compared to the true front. Removing the causal prior ($\boldsymbol{\omega} = \mathbf{0}$) increases the regret on all problems,

Table 7: Ablation experiment design of RESCUE. We sample $\widehat{\mathcal{D}}$ for causal discovery uniformly at random over $\mathbb{X} \times \mathbb{S}$, independently for each of the 20 seeds, with $|\widehat{\mathcal{D}}| = 500$ for Branin-Currin, Park, and Synthetic SCM and $|\widehat{\mathcal{D}}| = 1{,}000$ for robot navigation and constrained XGBoost HPO.

RESCUE variants	**Causal prior**	**Observational data**	**Reliability ω**
Causal prior contribution	×	Full	fixed to **0**
Observational data size	✓	A random subset of 10% and 50% of $\lvert\widehat{\mathcal{D}}\rvert$	Eq. (2)
Misspecified causal prior	✓	Outcomes of all samples shuffled across configuration–fidelity pairs	Eq. (2)

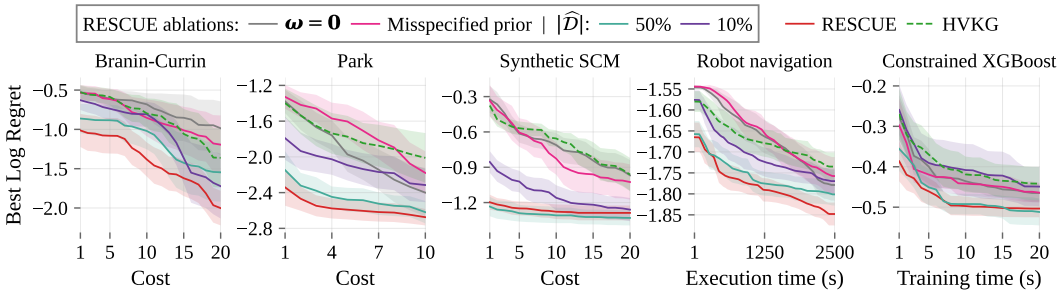


Figure 9: Cumulative intervention cost vs. best log HV regret (lower is better) for ablation experiments, where "best" denotes the minimum regret achieved up to that cost.

Table 8: Accuracy of the CPM proxy network distilled from *do*-calculus and its speedup over *do*-calculus. We compute the interventional mean $\bar{\boldsymbol{f}}_{\mathrm{do}}$ and standard deviation $\widehat{\boldsymbol{\sigma}}$ with *do*-calculus, which serves as the ground truth, and with the CPM proxy network at 500 held-out configuration–fidelity pairs. The RMSE of $\bar{\boldsymbol{f}}_{\mathrm{do}}$ is normalized by the standard deviation of the *do*-calculus values over the pairs and the RMSE of $\widehat{\boldsymbol{\sigma}}$ by their average.

Problem	Interventional mean $\bar{\boldsymbol{f}}_{\mathrm{do}}$		Interventional std. $\widehat{\boldsymbol{\sigma}}$	Time per query		Speedup
	R^2	Norm. RMSE	Norm. RMSE	*do*-calculus	CPM proxy	
Branin-Currin	0.986 ± 0.017	0.087 ± 0.051	0.060 ± 0.020	3.5 ms	0.19 μs	1.9×10^4
Park	0.996 ± 0.003	0.057 ± 0.025	0.024 ± 0.004	2.4 ms	0.21 μs	1.1×10^4
Synthetic SCM	1.000 ± 0.000	0.009 ± 0.002	0.023 ± 0.002	0.6 ms	0.19 μs	2.9×10^3
Robot navigation	0.930 ± 0.019	0.225 ± 0.028	0.057 ± 0.011	7.8 ms	0.20 μs	3.8×10^4
Constrained XGBoost HPO	0.929 ± 0.009	0.234 ± 0.013	0.068 ± 0.006	9.2 ms	0.19 μs	4.7×10^4

which shows that the improvement of RESCUE is contributed by the causal prior. The results also show that CPM remains informative even when learned from half of the observational data. The causal trust in Fig. 5 explains this behavior, where $\widehat{\omega}$ is lowest under the misspecified prior on all problems. This confirms that VoCI-HVKG correctly discounts the causal prior when it disagrees with the interventional data, and $\widehat{\omega}$ decreases with the size of $\widehat{\mathcal{D}}$. Fig. 10 compares the DAGs discovered under the causal prior of RESCUE and under the misspecified causal prior. Most edges of RESCUE are absent under this misspecified prior confirming the low $\widehat{\omega}$ reflects an incorrect causal prior. We also observe $\widehat{\omega}$ decreases over the optimization as interventional data accumulate demonstrating RESCUE's reliance on the causal prior only to the extent it is informative. Moreover, RESCUE remains on par with the strongest baseline HVKG on all problems even under the misspecified causal prior, which demonstrates its robustness to an incorrect causal prior.

A.6 Quality of causal model proxy

Table 8 reports the accuracy and runtime of the CPM proxy network against *do*-calculus. The proxy reproduces the interventional mean closely on every problem, and its error on the interventional

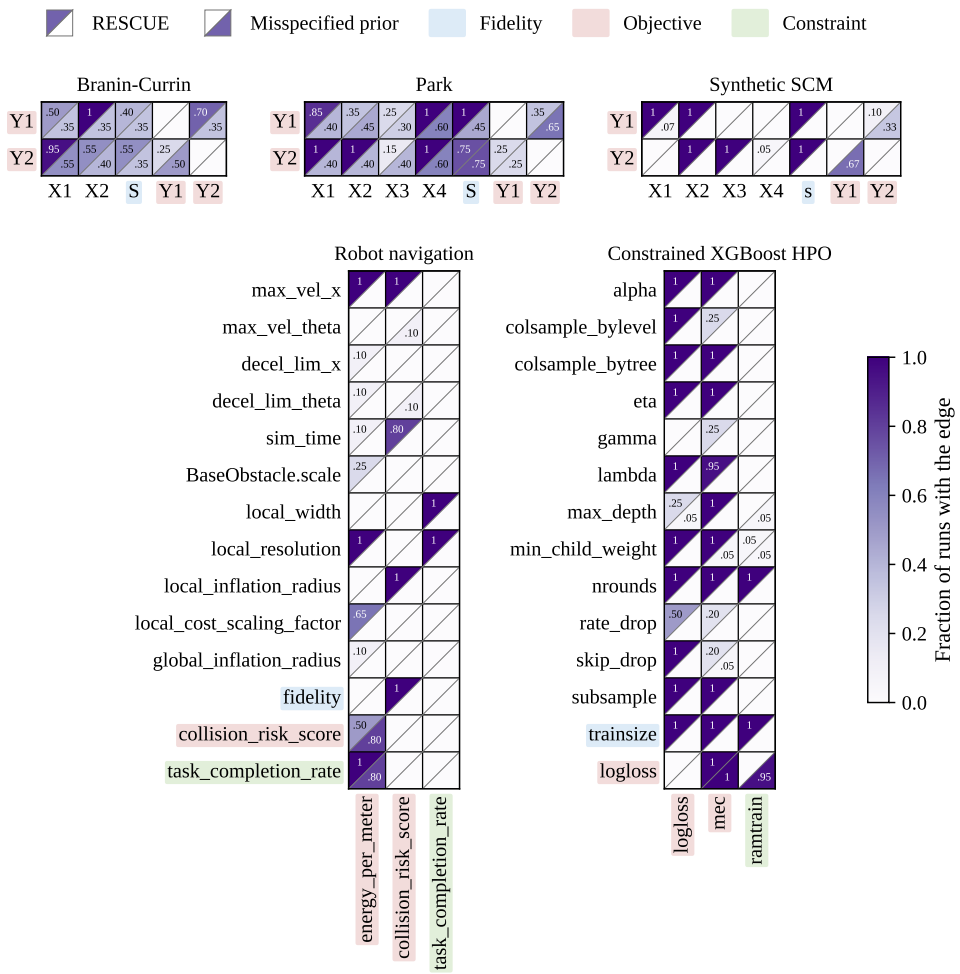


Figure 10: Impact of the misspecified causal prior of Table 7 on causal discovery. Each cell is the fraction of runs with the edge from the parent variable to the objective or constraint under the causal prior of RESCUE (upper-left triangle) and under the misspecified causal prior (lower-right triangle). For robot navigation, edges between $\mathbf{x}$ and edges in fewer than 5% of the runs are not shown.

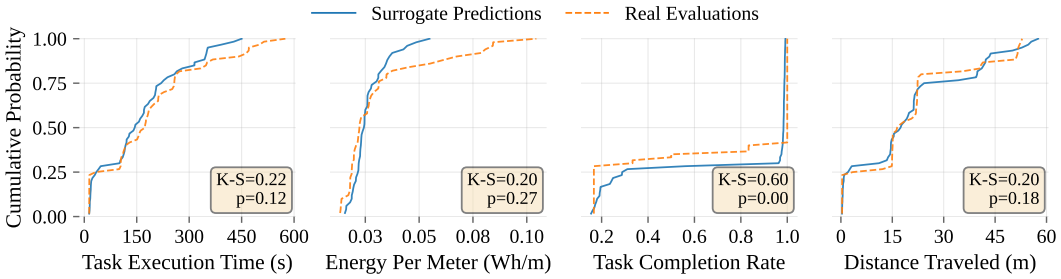
standard deviation is within the Monte Carlo noise of *do*-calculus. Averaged over the five problems, the proxy network achieves a speedup of 2.4×10^4. The causal prior is evaluated wherever the MF-CGP is evaluated, in the marginal likelihood during the MF-CGP fit, in the VoCI-HVKG acquisition, and in the NSGA-II search over the posterior mean. The acquisition alone requires $\approx 6 \times 10^4$ evaluations per iteration, at 512 initial candidates and 200 L-BFGS-B iterations over $1 + N_{\mathrm{f}} n_{\mathbb{P}}$ points each. These evaluations would require ≈ 0.6 to 9 minutes per iteration with *do*-calculus, while the proxy network requires ≈ 0.01 s. Therefore, the per-iteration runtime of RESCUE in Fig. 8 remains in the same order of magnitude as HVKG.

A.7 Multi-fidelity mobile robot navigation benchmark

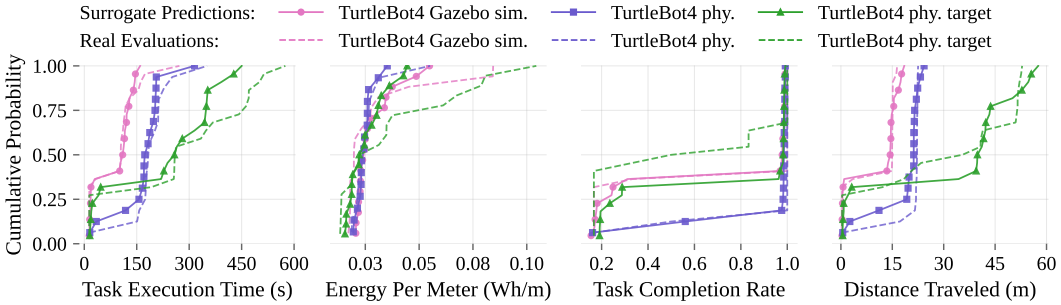
We construct a three-fidelity benchmark inspired by the setup of Hossen et al. (2025). Data are collected from: (i) *TurtleBot4 in Gazebo simulation* ($s = 0.2$), (ii) *TurtleBot4 physical robot in an obstacle-free indoor environment* ($s = 0.5$), and (iii) *TurtleBot4 physical robot in a complex indoor environment* (target, $s = 1.0$). We collect 100 interventional samples at each fidelity by uniformly sampling configurations within the parameter bounds listed in Table 10e, yielding 300 samples in total. For all trials, the robot is deployed under ROS 2, and evaluated on four task metrics: task completion rate, task execution time, energy per meter, and distance traveled. We randomly select

Table 9: Accuracy of the robot navigation benchmark model. Train and test R^2 and RMSE of each measured outcome over the 300 robot runs, 80% for training and 20% for testing.

Model output	Train		Test	
	R^2	RMSE	R^2	RMSE
Task execution time (s)	0.99	12.54	0.80	66.41
Energy per meter (Wh/m)	0.99	0.003	0.21	0.019
Task completion rate	1.00	0.01	0.68	0.21
Distance traveled (m)	0.99	1.58	0.75	8.07
Average	0.99	3.53	0.61	18.68



(a) Combined Empirical Cumulative Distribution Function (ECDF) for surrogate predictions and real evaluations.



(b) Per-fidelity ECDF for surrogate predictions and real evaluations.

Figure 11: Surrogate model validation for the robot navigation task. Close alignment between predicted and real distributions, combined with consistent performance across simulation and physical robot fidelities, demonstrates the surrogate's effectiveness.

80% of the data for training and reserve 20% for testing. $\widehat{\mathcal{D}}$ and every intervention of the optimization are evaluated on this model.

Surrogate Model Training and Architecture We fit independent CatBoost (Prokhorenkova et al., 2018) gradient-boosted regressors for task completion rate, task execution time, energy per meter, and distance traveled. Fidelity is included as a numeric covariate to retain its ordinal structure. The collision-risk score is computed from the configuration. All trained models are exported to ONNX format for cross-platform compatibility.

Distributional Alignment Across Fidelities Fig. 10a reports the Empirical Cumulative Distribution Function (ECDF) of surrogate predictions and real evaluations aggregated across fidelities. The Kolmogorov-Smirnov statistic (K-S and p-values) indicates close alignment for most metrics. The largest deviation is for task completion rate. Because this metric is bounded in $[0, 1]$, its empirical distribution often concentrates near 0 or 1 across all fidelities. Such boundary concentration reduces

variance and makes the K-S statistic more sensitive to small shifts between the surrogate and the real system, especially under the sample size of 100 per fidelity. Fig. 10b disaggregates ECDFs by fidelity and shows that the surrogate reproduces fidelity-dependent shifts, confirming it captures cross-fidelity trends and fidelity-specific variability.

Quantitative Performance Table 9 reports train and test R^2 and Root Mean Square Error (RMSE) for each metric. On the test set, the surrogate achieves an average R^2 of 0.61. Energy per meter has the smallest RMSE but the lowest R^2, distance traveled has a larger RMSE, and execution time has the largest RMSE. Combined with the distributional agreement in Fig. 11, the results indicate that the surrogate is sufficiently accurate to support the multi-fidelity mobile robot navigation benchmark, and any residual surrogate error propagates uniformly across all optimization methods including RESCUE and the baselines.

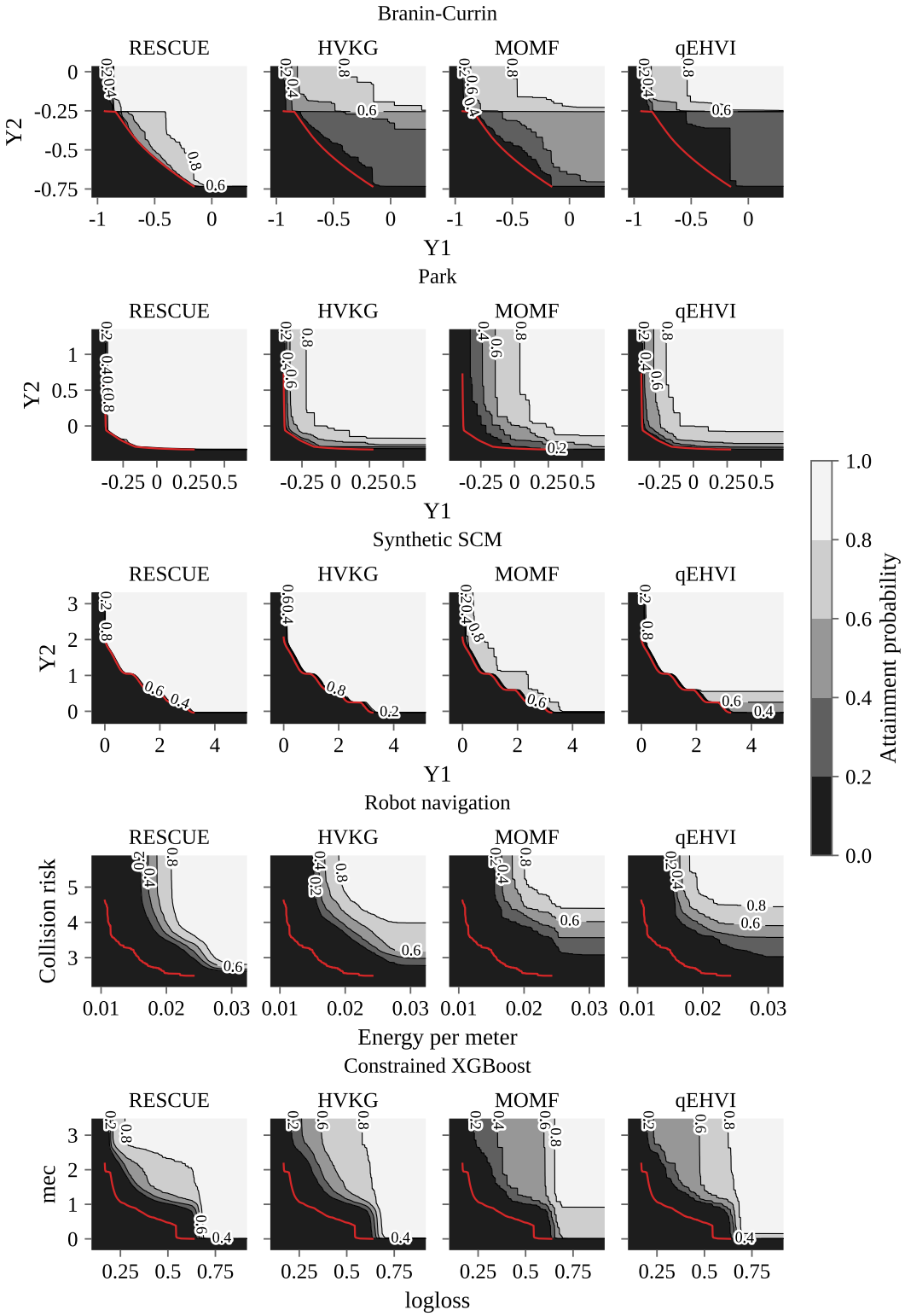


Figure 12: Empirical attainment function of the recommended Pareto set over the objective space. The true Pareto front is shown in red.

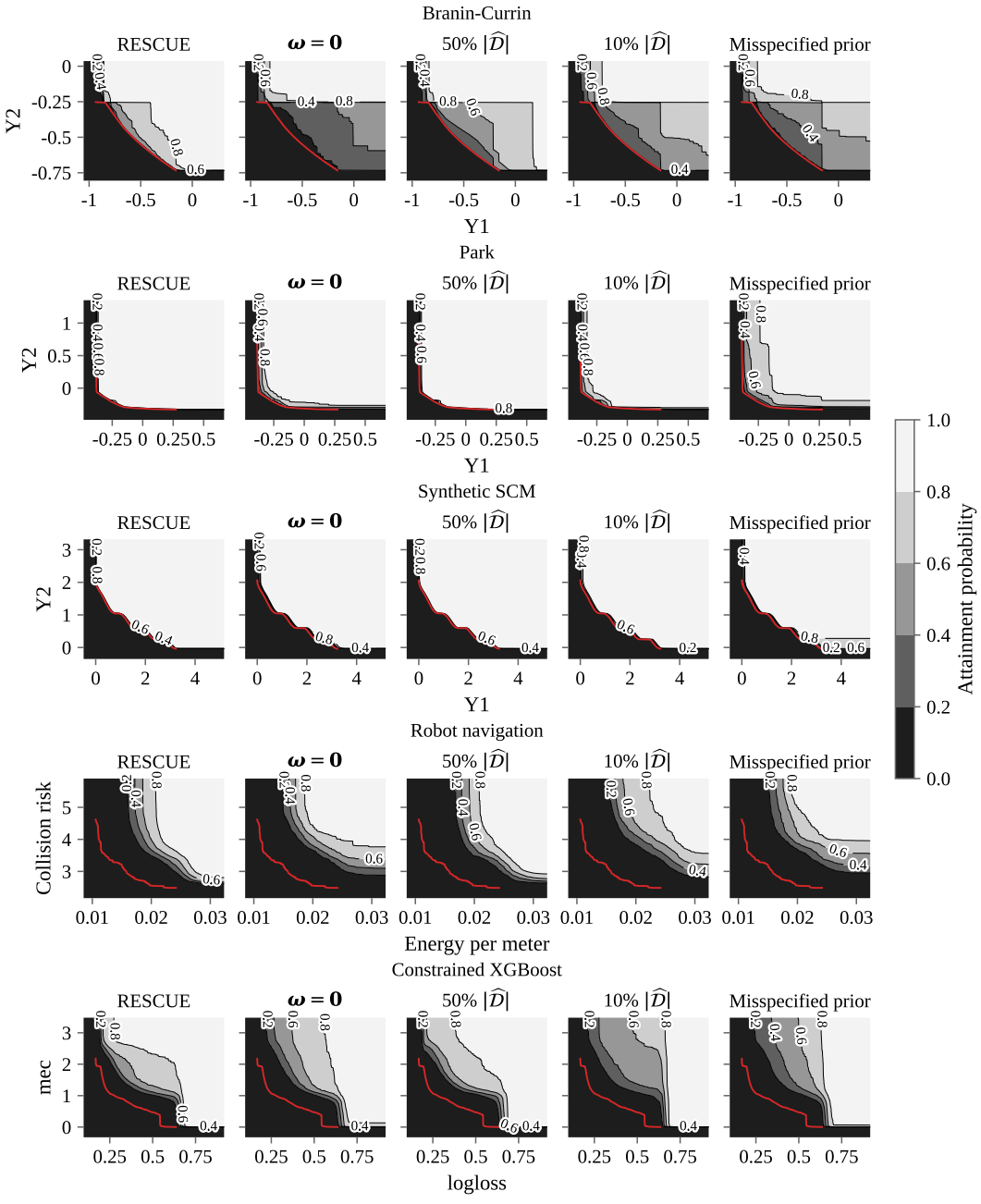


Figure 13: Empirical attainment function of the recommended Pareto set over the objective space for the ablation experiments. The true Pareto front is shown in red.

DISCOVERED DAGS AND CONFIGURATION SPACE

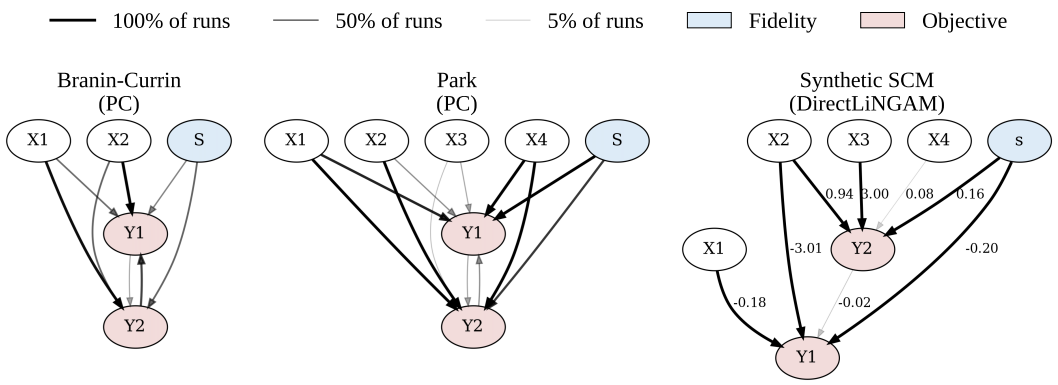


Figure 14: Discovered DAGs of RESCUE for Branin-Currin, Park and Synthetic SCM over 20 runs. Edge width and shade give the fraction of runs with the edge. Numbers are DirectLiNGAM coefficients averaged over the runs. Every evaluation intervenes on all $\mathbf{x}$, which removes the edges between them, so we exclude these edges and forbid edges from outputs to $\mathbf{x}$. Edges in fewer than 5% of the runs are not shown.



Figure 15: Discovered robot navigation DAG of RESCUE over 20 runs. Edge width and shade give the fraction of runs with the edge. We forbid edges from outputs to $\mathbf{x}$ and do not assume $\mathbf{x}$ are independent. Edges in fewer than 5% of the runs and options without an edge are not shown.

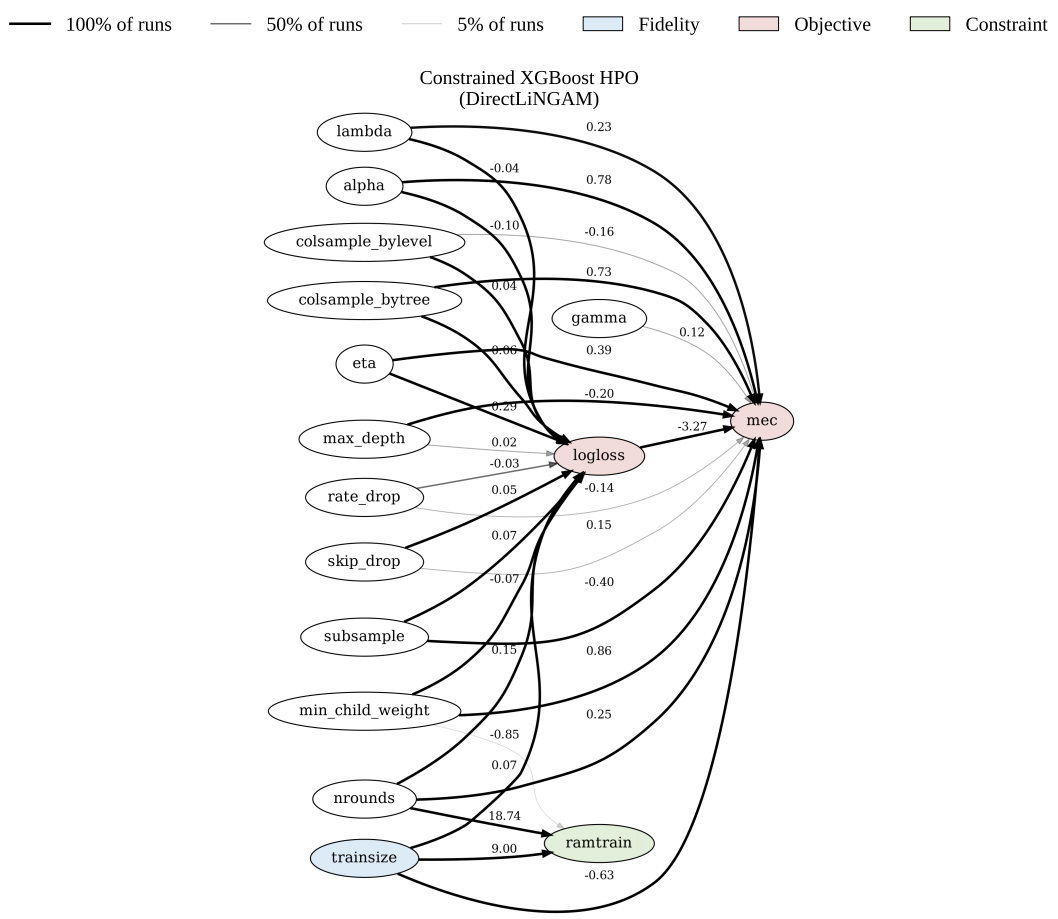


Figure 16: Discovered constrained XGBoost HPO DAG of RESCUE over 20 runs. Edge width and shade give the fraction of runs with the edge. Numbers are DirectLiNGAM coefficients averaged over the runs. Every evaluation intervenes on all $\mathbf{x}$, which removes the edges between them, so we exclude these edges and forbid edges from outputs to $\mathbf{x}$. Edges in fewer than 5% of the runs are not shown.

Table 10: Configuration Spaces of All Problems

(a) Branin–Currin

Configuration options	Range
X1	$[0, 1]$
X2	$[0, 1]$
S	$[0, 1]$

(b) Park

Configuration options	Range
X1	$[0, 1]$
X2	$[0, 1]$
X3	$[0, 1]$
X4	$[0, 1]$
S	$[0, 1]$

(c) Synthetic SCM

Configuration options	Range
X1	$[-2, 5]$
X2	$[0, 5]$
X3	$[0, 5]$
X4	$[-4, 5]$
S	$[0, 1]$

(d) Constrained XGBoost HPO (Scenario: iaml_xgboost, Instance: 41146)

Configuration options	Range/Value
alpha	$[10^{-4}, 10^{3}]$, log scale
colsample_bylevel	$[0.01, 1]$
colsample_bytree	$[0.01, 1]$
eta	$[10^{-4}, 1]$, log scale
gamma	$[10^{-4}, 7]$, log scale
lambda	$[10^{-4}, 10^{3}]$, log scale
max_depth	$[1, 15]$
min_child_weight	$[2.72, 150]$, log scale
nrounds	$[3, 2000]$, log scale
rate_drop	$[0, 1]$
skip_drop	$[0, 1]$
subsample	$[0.1, 1]$
trainsize (fidelity)	$[0.03, 1]$
booster	dart

(e) Robot Navigation Task

Configuration options	Range/Value
min_vel_x	$[-3, 0]$
max_vel_x	$[0.1, 0.5]$
max_vel_theta	$[0.5, 2.0]$
max_speed_xy	$[0, 0.1]$
acc_lim_x	$[1.5, 4]$
acc_lim_theta	$[1.5, 5]$
decel_lim_x	$[-4, -1]$
decel_lim_theta	$[-5, -1.5]$
vx_samples	$[10, 40]$
vtheta_samples	$[10, 40]$
sim_time	$[1, 3]$
BaseObstacle.scale	$[0.01, 0.1]$
PathAlign.scale	$[10, 60]$
GoalAlign.scale	$[10, 60]$
PathDist.scale	$[10, 60]$
GoalDist.scale	$[10, 60]$
RotateToGoal.scale	$[10, 60]$
local_width	$[2, 5]$
local_height	$[2, 5]$
local_resolution	$[0.04, 0.1]$
local_inflation_radius	$[0.3, 0.6]$
local_cost_scaling_factor	$[2, 10]$
global_resolution	$[0.04, 0.1]$
global_inflation_radius	$[0.3, 0.6]$
global_cost_scaling_factor	$[2, 10]$
xy_goal_tolerance	0.25
yaw_goal_tolerance	0.25
fidelity	$0.2, 0.5, 1.0$

B PROOFS

Throughout, we adopt the extended-domain view of section 5, in which the MF-CGP is a scalar GP over triples of configuration, fidelity, and output index, with kernel $\tilde{k} \leq 1$, prior mean $\bar{\boldsymbol{f}}_{\mathrm{do}}$, and noise variance σ^2. Each intervention contributes the J scalar observations $\{((\boldsymbol{x}_t, s_t), j)\}_{j=1}^{J}$. We write $\zeta = ((\boldsymbol{x}, s), j)$ for a generic triple, $\mathbb{Z}_t$ for the set of the Jt triples observed up to intervention t, and $\gamma_t := \max_{\boldsymbol{\chi}_1,\ldots,\boldsymbol{\chi}_t} I((\mathbf{y}, \mathbf{h})_{1:t}; \bar{\boldsymbol{f}})$ for the maximum information gain after t vector-valued interventions. As is standard in GP-based BO analyses, we fix the kernel hyperparameters throughout, although, in practice, they are selected by maximizing the log marginal likelihood. The analysis depends on the kernel only through Assumption 5.1 and the information gain. Therefore, the analysis applies to any bounded kernel. By Assumption 5.2, $w \in \mathbb{H}_{\tilde{k}}$ with $\|w\|_{\mathbb{H}_{\tilde{k}}} \leq B$ such that, writing

$$\boldsymbol{\Delta} := \bar{\boldsymbol{f}} - \bar{\boldsymbol{f}}_{\mathrm{do}} - \boldsymbol{w}, \qquad \|\boldsymbol{\Delta}\|_\infty \leq \xi.$$

B.1 PROOF OF LEMMA 5.6

Lemma 5.6 is important for translating a bound on the posterior prediction error into a bound on the hypervolume regret of the recommended set. In particular, it establishes that the hypervolume indicator is Lipschitz-continuous under uniform perturbations of the objective values, with the explicit constant $L = 2MR^{M-1}$, without requiring a separate assumption on the hypervolume indicator. We prove the result in three steps. First, we bound the hypervolume change when one objective changes by a fixed amount. We then extend the bound to equal changes in all objectives. Finally, we use these bounds to control the hypervolume change under any perturbation of at most ϵ per objective.

Proof. Let $\boldsymbol{a} : \mathbb{X}_0 \to [\boldsymbol{r} - R\mathbf{1}, \boldsymbol{r}]$. Define the dominated region as $\mathbb{O}(\boldsymbol{a}) := \{\boldsymbol{z} \in [\boldsymbol{r} - R\mathbf{1}, \boldsymbol{r}] : \exists\, \boldsymbol{x} \in \mathbb{X}_0,\ \boldsymbol{a}(\boldsymbol{x}) \preceq \boldsymbol{z} \preceq \boldsymbol{r}\}$. Its volume gives the hypervolume $\mathrm{HV}[\{\boldsymbol{a}(\boldsymbol{x})\}_{\boldsymbol{x}\in\mathbb{X}_0}] = \lambda(\mathbb{O}(\boldsymbol{a}))$ where λ denotes Lebesgue measure. Since $\mathbb{O}(\boldsymbol{a})$ is an upper set, each nonempty slice with $\boldsymbol{z}_{-m}$ fixed is an interval with lower boundary $\ell_m(\boldsymbol{z}_{-m})$.

Step 1 (single-axis shift). Consider $\boldsymbol{a} - \epsilon\boldsymbol{e}^{(m)}$ for $\epsilon \geq 0$ with values projected onto the box. This subtracts ϵ from objective m only. Let $\ell_m^-(\boldsymbol{z}_{-m})$ denote the corresponding lower boundary of the dominated region. This boundary decreases by at most ϵ, giving

$$\lambda\big(\mathbb{O}(\boldsymbol{a} - \epsilon\boldsymbol{e}^{(m)})\big) - \lambda\big(\mathbb{O}(\boldsymbol{a})\big) = \int \big[\ell_m(\boldsymbol{z}_{-m}) - \ell_m^-(\boldsymbol{z}_{-m})\big]\, d\boldsymbol{z}_{-m} \leq \epsilon\, R^{M-1},$$

where $\boldsymbol{z}_{-m}$ lies in an $(M-1)$-dimensional box of volume R^{M-1}.

Step 2 (diagonal shift). Writing $\boldsymbol{a} - \epsilon\mathbf{1}$ as a composition of M single-axis shifts and summing the bound from Step 1 over the M axes gives $\lambda\big(\mathbb{O}(\boldsymbol{a} - \epsilon\mathbf{1})\big) - \lambda\big(\mathbb{O}(\boldsymbol{a})\big) \leq M\epsilon R^{M-1}$.

Step 3 (arbitrary perturbations). If $\sup_{\boldsymbol{x}\in\mathbb{X}_0} \|\boldsymbol{a}(\boldsymbol{x}) - \boldsymbol{b}(\boldsymbol{x})\|_\infty \leq \epsilon$, then $\boldsymbol{a}(\boldsymbol{x}) - \epsilon\mathbf{1} \preceq \boldsymbol{b}(\boldsymbol{x}) \preceq \boldsymbol{a}(\boldsymbol{x}) + \epsilon\mathbf{1}$ for every $\boldsymbol{x} \in \mathbb{X}_0$. Therefore $\mathbb{O}(\boldsymbol{a} + \epsilon\mathbf{1}) \subseteq \mathbb{O}(\boldsymbol{b}) \subseteq \mathbb{O}(\boldsymbol{a} - \epsilon\mathbf{1})$. Applying the bound from Step 2 twice gives

$$\big|\mathrm{HV}[\{\boldsymbol{a}(\boldsymbol{x})\}] - \mathrm{HV}[\{\boldsymbol{b}(\boldsymbol{x})\}]\big| \leq \lambda\big(\mathbb{O}(\boldsymbol{a} - \epsilon\mathbf{1})\big) - \lambda\big(\mathbb{O}(\boldsymbol{a} + \epsilon\mathbf{1})\big) \leq 2MR^{M-1}\epsilon = L\,\epsilon,$$

which completes the proof. □

B.2 PROOF OF LEMMA 5.7

The inclusion of Lemma 5.7 is critical for establishing the theoretical validity of using *estimated* causal priors within a Gaussian Process bandit framework. Standard regret analyses typically rely on a strong "realizable" assumption that the true objective function lies exactly within the Reproducing Kernel Hilbert Space (RKHS) associated with the GP prior. In RESCUE, however, the causal prior is derived from observational data and is an *approximation*, not a perfect oracle. By adopting the agnostic setting of Assumption 5.2, Lemma 5.7 rigorously quantifies how the causal misspecification error ξ propagates into the posterior predictions.

Moreover, this result provides the key guarantee underpinning the behavior of the algorithm under model mismatch. As long as the causal prior is a reasonable approximation (i.e., ξ is bounded), the

true objective function remains within the inflated confidence intervals with high probability. This directly connects our causal inference component to the robust GP bandit literature (e.g., Bogunovic & Krause (2021)), and shows that RESCUE does not require a perfectly specified causal model to succeed, instead degrades gracefully as a function of the causal estimation error.

Proof. We work in the standard agnostic or misspecified GP setting of Assumption 5.2. The GP posterior update with prior mean $\bar{\boldsymbol{f}}_{\mathrm{do}}$ and noise variance σ^2 is

$$\mu_t(\zeta) - \bar{\boldsymbol{f}}_{\mathrm{do}}(\zeta) = \boldsymbol{k}_*(\zeta)^\top (\boldsymbol{\Sigma}_t + \sigma^2 \boldsymbol{I})^{-1} \big(\boldsymbol{w}(\mathbb{Z}_t) + \boldsymbol{\Delta}(\mathbb{Z}_t) + \boldsymbol{\epsilon}_{1:t}\big),$$

where $\boldsymbol{k}_*(\zeta)$ and $\boldsymbol{\Sigma}_t$ denote the kernel vector and Gram matrix over $\mathbb{Z}_t$ defined in Section 4.1. The vector $\boldsymbol{\epsilon}_{1:t}$ denotes the stacked observation noise. We decompose the prediction error as

$$\bar{f}(\zeta) - \mu_t(\zeta) = \underbrace{\Big[w(\zeta) - \boldsymbol{k}_*(\zeta)^\top (\boldsymbol{\Sigma}_t + \sigma^2 \boldsymbol{I})^{-1} \big(\boldsymbol{w}(\mathbb{Z}_t) + \boldsymbol{\epsilon}_{1:t}\big)\Big]}_{\text{(I)}} + \underbrace{\Delta(\zeta)}_{\text{(II)}} - \underbrace{\boldsymbol{k}_*(\zeta)^\top (\boldsymbol{\Sigma}_t + \sigma^2 \boldsymbol{I})^{-1} \boldsymbol{\Delta}(\mathbb{Z}_t)}_{\text{(III)}}.$$

Term (I) is the GP prediction error for the RKHS function w with $\|w\| \le B$. We apply the self-normalized confidence bound of Chowdhury & Gopalan (2017, Theorem 2) (see also Srinivas et al. 2009) using regularization σ^2 and σ-sub-Gaussian noise. The noise coefficient simplifies to $\sigma/\sqrt{\sigma^2} = 1$. For any $\rho \in (0, 1)$, the following bound holds with probability at least $1 - \rho$ uniformly over all ζ and $t \ge 1$.

$$|(\mathrm{I})| \le \Big(B + \sqrt{2\big(\gamma_t + 1 + \ln(1/\rho)\big)}\Big)\, \sigma_t(\zeta) = \beta_t^{1/2} \sigma_t(\zeta),$$

where the information gain from the t interventions is at most γ_t. Since ζ includes the output index j, the bound is also uniform over the outputs, and no separate union bound is needed.

For the second term, $|(\mathrm{II})| \le \xi$ by Assumption 5.2.

To bound term (III), let $\boldsymbol{\phi}$ denote the feature map of $\tilde{k}$ and $\boldsymbol{\Phi}_t$ the feature matrix over $\mathbb{Z}_t$. Define $\boldsymbol{A}_t := \boldsymbol{\Phi}_t^\top \boldsymbol{\Phi}_t + \sigma^2 \boldsymbol{I}$. Then $(\mathrm{III}) = \boldsymbol{\phi}(\zeta)^\top \boldsymbol{A}_t^{-1} \boldsymbol{\Phi}_t^\top \boldsymbol{\Delta}(\mathbb{Z}_t)$. Applying Cauchy–Schwarz in the $\boldsymbol{A}_t^{-1}$-norm gives

$$|(\mathrm{III})| \le \|\boldsymbol{\phi}(\zeta)\|_{\boldsymbol{A}_t^{-1}} \cdot \|\boldsymbol{\Phi}_t^\top \boldsymbol{\Delta}(\mathbb{Z}_t)\|_{\boldsymbol{A}_t^{-1}}.$$

The posterior-variance identity $\sigma_t^2(\zeta) = \sigma^2 \boldsymbol{\phi}(\zeta)^\top \boldsymbol{A}_t^{-1} \boldsymbol{\phi}(\zeta)$ gives $\|\boldsymbol{\phi}(\zeta)\|_{\boldsymbol{A}_t^{-1}} = \sigma_t(\zeta)/\sigma$. For $\|\boldsymbol{\Phi}_t^\top \boldsymbol{\Delta}(\mathbb{Z}_t)\|_{\boldsymbol{A}_t^{-1}}$ we have

$$\|\boldsymbol{\Phi}_t^\top \boldsymbol{\Delta}(\mathbb{Z}_t)\|_{\boldsymbol{A}_t^{-1}}^2 = \boldsymbol{\Delta}(\mathbb{Z}_t)^\top \boldsymbol{\Phi}_t \boldsymbol{A}_t^{-1} \boldsymbol{\Phi}_t^\top \boldsymbol{\Delta}(\mathbb{Z}_t) \le \|\boldsymbol{\Delta}(\mathbb{Z}_t)\|_2^2 \le \xi^2\, Jt,$$

where $\lambda_{\max}(\boldsymbol{\Phi}_t \boldsymbol{A}_t^{-1} \boldsymbol{\Phi}_t^\top) \le 1$ and $|\mathbb{Z}_t| = Jt$. Combining these bounds gives

$$|\bar{f}(\zeta) - \mu_t(\zeta)| \le |(\mathrm{I})| + |(\mathrm{II})| + |(\mathrm{III})| \le \beta_t^{1/2} \sigma_t(\zeta) + \xi\Big(1 + \tfrac{1}{\sigma}\sqrt{Jt}\, \sigma_t(\zeta)\Big),$$

which is the claimed inequality and completes the proof. □

B.3 A FEASIBILITY SANDWICH

The recommendation extracts the Pareto set over the estimated feasible region $\widehat{\mathbb{X}}_T$, which can differ from $\bar{\mathbb{X}}$ in two ways. $\widehat{\mathbb{X}}_T$ can admit a configuration that is not feasible, and it can omit a feasible configuration whose objectives contribute to HV^*. Lemma B.1 bounds constraint violations within $\widehat{\mathbb{X}}_T$ in proportion to the causal-model error ξ. In the well-specified setting this gives $\widehat{\mathbb{X}}_T \subseteq \bar{\mathbb{X}}$. The lemma also establishes $\bar{\mathbb{X}}_{2\bar{\eta}} \subseteq \widehat{\mathbb{X}}_T$. The resulting hypervolume gap is at most $\mathrm{HV}^* - \mathrm{HV}^*_{2\bar{\eta}}$, which Assumption 5.5 bounds by $2L'\bar{\eta}$.

We work on the high-probability event of Lemma 5.7 at intervention T and write, for each constraint q and $\boldsymbol{x} \in \mathbb{X}$,

$$\eta_q(\boldsymbol{x}) := \beta_T^{1/2} \sigma_{T, M+q}(\boldsymbol{x}, s_\diamond) + \xi\Big(1 + \tfrac{1}{\sigma}\sqrt{JT}\, \sigma_{T, M+q}(\boldsymbol{x}, s_\diamond)\Big), \qquad \bar{\eta} := \sup_{\boldsymbol{x}} \max_q \eta_q(\boldsymbol{x}),$$

so that $|h_q(\boldsymbol{x}, s_\diamond) - \mu_{T, M+q}(\boldsymbol{x}, s_\diamond)| \le \eta_q(\boldsymbol{x})$ for all $\boldsymbol{x}, q$.

Lemma B.1 (Feasibility sandwich)**.** *On the event of Lemma 5.7, the region* $\widehat{\mathbb{X}}_T = \{\boldsymbol{x} : \mu_{T,M+q} \geq \beta_T^{1/2}\sigma_{T,M+q}\ \forall q\}$ *satisfies: (i) every* $\boldsymbol{x} \in \widehat{\mathbb{X}}_T$ *satisfies* $h_q(\boldsymbol{x}, s_\diamond) \geq -\xi\big(1 + \frac{1}{\sigma}\sqrt{JT}\,\sup_{\boldsymbol{x}'}\max_{q'}\sigma_{T,M+q'}(\boldsymbol{x}', s_\diamond)\big)$ *for all q, with* $\widehat{\mathbb{X}}_T \subseteq \bar{\mathbb{X}}$ *when* $\xi = 0$*; (ii)* $\widehat{\mathbb{X}}_T \supseteq \bar{\mathbb{X}}_{2\bar{\eta}}$.

Proof. For part (i), let $\boldsymbol{x} \in \widehat{\mathbb{X}}_T$. By the definition of $\widehat{\mathbb{X}}_T$ and the confidence bound above, $h_q \geq \mu_{T,M+q} - \eta_q(\boldsymbol{x}) \geq \beta_T^{1/2}\sigma_{T,M+q} - \eta_q(\boldsymbol{x}) = -\xi\big(1 + \frac{1}{\sigma}\sqrt{JT}\sigma_{T,M+q}\big)$.

For part (ii), suppose that $h_q(\boldsymbol{x}, s_\diamond) \geq 2\bar{\eta}$ for all q. Then $\mu_{T,M+q} \geq h_q - \eta_q(\boldsymbol{x}) \geq 2\bar{\eta} - \eta_q(\boldsymbol{x}) \geq \bar{\eta} \geq \beta_T^{1/2}\sigma_{T,M+q}$, where the last inequality follows from $\beta_T^{1/2}\sigma_{T,M+q} \leq \eta_q(\boldsymbol{x}) \leq \bar{\eta}$. Thus, $\boldsymbol{x} \in \widehat{\mathbb{X}}_T$, which completes the proof. □

B.4 Proof of Theorem 5.8

Proof. By Lemma 5.7, with probability at least $1 - \rho$, the confidence bound holds uniformly over all configurations and outputs. Let

$$\bar{\sigma}_T := \sup_{\boldsymbol{x}} \|\boldsymbol{\sigma}_T(\boldsymbol{x}, s_\diamond)\|_2, \qquad \Gamma_T := \beta_T^{1/2}\bar{\sigma}_T + \xi\Big(1 + \tfrac{1}{\sigma}\sqrt{JT}\,\bar{\sigma}_T\Big).$$

Step 1 (uniform prediction error). For any $\boldsymbol{x}$ and output j, the j-th component of $\boldsymbol{\sigma}_T$ is bounded by its Euclidean norm, $\sigma_{T,j}(\boldsymbol{x}, s_\diamond) \leq \|\boldsymbol{\sigma}_T(\boldsymbol{x}, s_\diamond)\|_2 \leq \bar{\sigma}_T$. Applying Lemma 5.7 gives

$$\sup_{\boldsymbol{x}} \max_{j \leq J} \big|\bar{f}_j(\boldsymbol{x}, s_\diamond) - \mu_{T,j}(\boldsymbol{x}, s_\diamond)\big| \;\leq\; \Gamma_T.$$

For the objective outputs, the prediction error is bounded by Γ_T. For the constraint outputs, the error bound $\bar{\eta}$ used in Lemma B.1 satisfies $\bar{\eta} \leq \Gamma_T$.

Step 2 (feasibility of the recommendation). By Lemma B.1(i) and $\sigma_{T,M+q}(\boldsymbol{x}, s_\diamond) \leq \bar{\sigma}_T$, every $\boldsymbol{x} \in \widehat{\mathbb{P}}_T$ satisfies, for all q,

$$h_q(\boldsymbol{x}, s_\diamond) \;\geq\; -\xi\Big(1 + \tfrac{1}{\sigma}\sqrt{JT}\,\bar{\sigma}_T\Big).$$

Step 3 (projection onto the box). The recommendation uses the objective posterior means projected onto the box of Assumption 5.3. Projecting each objective posterior mean onto its allowed interval cannot increase its error because the true objective value lies within that interval. Thus, Lemma 5.6 applies to the projected posterior-mean objective vectors and the true objective vectors.

Step 4 (combine). Let $\mathrm{HV}_{\boldsymbol{\mu}}[\cdot]$ and $\mathrm{HV}_{\boldsymbol{f}}[\cdot]$ denote the hypervolume of the posterior-mean and true objective values, respectively, of a set of configurations at $s_\diamond$. Combining the bounds above gives

$$\begin{aligned}
\mathrm{HV}_{\boldsymbol{f}}[\widehat{\mathbb{P}}_T] \;&\geq\; \mathrm{HV}_{\boldsymbol{\mu}}[\widehat{\mathbb{P}}_T] - L\,\Gamma_T && \text{(Lemma 5.6 with } \mathbb{X}_0 = \widehat{\mathbb{P}}_T\text{)}\\
&=\; \mathrm{HV}_{\boldsymbol{\mu}}[\widehat{\mathbb{X}}_T] - L\,\Gamma_T && (\widehat{\mathbb{P}}_T \text{ is the } \boldsymbol{\mu}_T\text{-Pareto set of } \widehat{\mathbb{X}}_T)\\
&\geq\; \mathrm{HV}_{\boldsymbol{\mu}}[\bar{\mathbb{X}}_{2\bar{\eta}}] - L\,\Gamma_T && \text{(Lemma B.1(ii), monotonicity of HV)}\\
&\geq\; \mathrm{HV}_{\boldsymbol{f}}[\bar{\mathbb{X}}_{2\bar{\eta}}] - 2L\,\Gamma_T && \text{(Lemma 5.6 with } \mathbb{X}_0 = \bar{\mathbb{X}}_{2\bar{\eta}}\text{)}\\
&=\; \mathrm{HV}^*_{2\bar{\eta}} - 2L\,\Gamma_T \;\geq\; \mathrm{HV}^* - 2L'\bar{\eta} - 2L\,\Gamma_T && \text{(Assumption 5.5).}
\end{aligned}$$

Finally, using $\bar{\eta} \leq \Gamma_T$ yields $\mathrm{HV}^* - \mathrm{HV}_{\boldsymbol{f}}[\widehat{\mathbb{P}}_T] \;\leq\; 2(L + L')\,\Gamma_T$, which is exactly Eq. (5) and completes the proof. □

Remark B.2 (Reliability mixture). If Assumption 5.2 holds for both scaled priors $\boldsymbol{\omega} \odot \bar{\boldsymbol{f}}_{\mathrm{do}} + (\mathbf{1} - \boldsymbol{\omega}) \odot \boldsymbol{c}$, $\boldsymbol{\omega} \in \{\mathbf{0}, \mathbf{1}\}$, with common constants (B, ξ), then Theorem 5.8 holds for $\widehat{\mathbb{P}}_T$ computed from the mixture posterior mean over $\widehat{\mathbb{X}}_T$, where $\ln(2/\rho)$ replaces $\ln(1/\rho)$ in β_T and $\bar{\sigma}_T = \sup_{\boldsymbol{x}}\big(\sum_j \max_{a\in\{0,1\}} \sigma^2_{T,j,a}(\boldsymbol{x}, s_\diamond)\big)^{1/2}$ with $\sigma_{T,j,a}$ the posterior standard deviation of output j under $\mathcal{M}_{\boldsymbol{a}}$. A union bound ensures that both models' confidence bounds hold simultaneously with probability at least $1 - \rho$. Each mixture mean is a convex combination of the two model means, so the same error bounds apply using the maximum component uncertainty. The proof of Theorem 5.8 then applies to the mixture recommendation.

B.5 Proof of Corollary 5.9

Proof. Part (i). By Theorem 5.8, with probability at least $1-\rho$, the hypervolume regret of $\widehat{\mathbb{P}}_T$ is bounded by

$$2\big(L+L'\big)\left(\beta_T^{1/2}\,\bar{\sigma}_T \;+\; \xi\Big(1+\tfrac{1}{\sigma}\sqrt{JT}\,\bar{\sigma}_T\Big)\right).$$

The constants $L = 2MR^{M-1}$ and L' depend only on the number of objectives, the box of Assumption 5.3, and the margin regularity of Assumption 5.5. The confidence parameter $\beta_T^{1/2} = B + \sqrt{2(\gamma_T + 1 + \ln(1/\rho))}$ depends on the residual RKHS bound B of Assumption 5.2 and the information gain γ_T, which is a property of the kernel. By Assumption 5.2, the constants B and ξ are defined with respect to the causal prior $\bar{\boldsymbol{f}}_{\mathrm{do}}$, which is identified by do-calculus from the observational data $\widehat{\mathcal{D}}$. For a family of auxiliary sources satisfying Assumption 5.2 with common (B, ξ), these constants do not depend on the cross-fidelity bias $\boldsymbol{\delta}(\boldsymbol{x}, s)$.

The remaining quantity, $\bar{\sigma}_T = \sup_{\boldsymbol{x}} \|\boldsymbol{\sigma}_T(\boldsymbol{x}, s_\diamond)\|_2$, depends on the MF-CGP posterior variance at the target fidelity. For fixed kernel hyperparameters, this variance is independent of the observed function values and is determined only by the kernel, the query locations, and the noise level. Although cross-fidelity bias $\boldsymbol{\delta}(\boldsymbol{x}, s)$ may affect the auxiliary observations, it does not affect $\boldsymbol{\sigma}_T(\boldsymbol{x}, s_\diamond)$ for a fixed realized query sequence. Thus the regret bound has no explicit dependence on the magnitude of this bias. Consequently, the regret bound above holds with the same constants for any such family with $\sup_{\boldsymbol{x}, s \neq s_\diamond} \|\boldsymbol{\delta}(\boldsymbol{x}, s)\| \to \infty$. The bias may affect query selection and the rate at which $\bar{\sigma}_T$ decreases, but the bound remains valid for the realized query sequence. This establishes the claimed robustness to unreliable auxiliary sources.

Part (ii). Restrict the recommendation posterior to the causal prior and the target-fidelity observations by conditioning the GP only on interventions performed at $s_\diamond$. The restricted model remains a GP posterior over the slice $\mathbb{X} \times \{s_\diamond\}$, with the restricted kernel, its own information gain, and the number of target-fidelity observations in place of T. Lemma 5.7 applies directly to this posterior. The feasibility bounds of Lemma B.1 and the hypervolume inequalities in the proof of Theorem 5.8 remain valid. Therefore, Theorem 5.8 holds with the posterior mean, posterior variance, information gain, and realized uncertainty $\bar{\sigma}_T$ defined on the target-fidelity slice. Since this posterior is conditioned only on target-fidelity observations, Assumption 5.2 is required only at $s_\diamond$. This completes the proof. □

B.6 A sublinear rate under periodic target-fidelity exploration

Theorem 5.8 bounds the regret by the realized uncertainty $\bar{\sigma}_T$, and this section quantifies how fast $\bar{\sigma}_T$ can be driven to zero. Algorithm 3 is a variant of Algorithm 1 with exploration cycle N_{e}. At every N_{e}-th intervention, the algorithm queries the target fidelity at the point of maximum aggregated posterior uncertainty (Line 7). At all remaining interventions, it optimizes the VoCI-HVKG acquisition in Eq. (4) (Line 9). Thus, the exploration query replaces acquisition optimization only at the scheduled exploration interventions. Setting $N_{\mathrm{e}} = \infty$ recovers pure VoCI-HVKG. The remaining budget must cover the target-fidelity query scheduled at every N_{e}-th intervention. Under the normalized costs of Assumption 5.4, this requires $\Lambda \geq \lfloor T/N_{\mathrm{e}} \rfloor$. If this condition is not satisfied, the anytime guarantee of Theorem 5.8 remains valid, but the convergence rate derived below does not apply.

Lemma B.3 (Uniform variance decay through exploration)**.** *Fix the kernel hyperparameters. Under Assumption 5.1, after $T \geq \max\{2N_0, 2N_{\mathrm{e}}\}$ interventions of RESCUE with exploration cycle N_{e},*

$$\sup_{\boldsymbol{x} \in \mathbb{X}} \|\boldsymbol{\sigma}_T(\boldsymbol{x}, s_\diamond)\|_2^2 \;\leq\; \frac{4\,K\,N_{\mathrm{e}}\,\gamma_T}{T}, \qquad K := \frac{2J}{\ln\big(1 + J\,\sigma^{-2}\big)}.$$

The proof uses three properties of the MF-CGP posterior. First, the posterior variance does not increase as data accumulate. Second, each exploration query selects the point with the largest uncertainty at the target fidelity. Third, the information gain of the kernel bounds the total posterior variance accumulated along the query sequence. Together, these properties bound the largest remaining uncertainty by the average uncertainty at the exploration queries.

Proof. Let $t_1, \ldots, t_n$ denote the exploration interventions, which occur at every N_{e}-th iteration of the optimization loop. The loop performs $T - N_0$ interventions, so $n = \lfloor (T - N_0)/N_{\mathrm{e}} \rfloor \geq T/(4N_{\mathrm{e}})$

Algorithm 3 RESCUE with Periodic Target-Fidelity Exploration

Require: $\widehat{\mathcal{D}}$, $\mathbb{X}$, $\mathbb{S}$, $s_\diamond$, $c(\cdot,\cdot)$, budget Λ, exploration cycle N_e
Ensure: Recommended Pareto set
1: Initialize $\mathcal{D}_0 = \{(\boldsymbol{x}_i, s_i, \boldsymbol{y}_i)\}_{i=1}^{N_0}$ (Appendix A.1), $t = 0$, and $C_0 = 0$
2: Construct a causal performance model (CPM): $\mathcal{C}(\mathcal{G} \mid \widehat{\mathcal{D}})$
3: **while** $C_t \leq \Lambda$ **do**
4: $\quad t \leftarrow t + 1$
5: $\quad \mathcal{M}(\boldsymbol{x}, s \mid \mathcal{D}_{0:t-1}, \mathcal{C}) \leftarrow$ Train a MF-CGP
6: $\quad$ **if** $t \bmod N_\mathrm{e} = 0$ **then**
7: $\quad\quad s_t \leftarrow s_\diamond$ and $\boldsymbol{x}_t \leftarrow \arg\max_{\boldsymbol{x}\in\mathbb{X}} \|\boldsymbol{\sigma}_{t-1}(\boldsymbol{x}, s_\diamond)\|_2$
8: $\quad$ **else**
9: $\quad\quad$ Find next $(\boldsymbol{x}_t, s_t)$ by optimizing Eq. (4)
10: $\quad$ **end if**
11: $\quad$ **if** $C_{t-1} + c(\boldsymbol{x}_t, s_t) > \Lambda$ **then**
12: $\quad\quad$ **break**
13: $\quad$ **end if**
14: $\quad$ Obtain $\boldsymbol{y}_t$ by intervening $\boldsymbol{x}_t$ at s_t
15: $\quad$ Augment $\mathcal{D}_{0:t} \leftarrow \mathcal{D}_{0:t-1} \cup \{(\boldsymbol{x}_t, s_t, \boldsymbol{y}_t)\}$
16: $\quad$ Update total cost $C_t \leftarrow C_{t-1} + c(\boldsymbol{x}_t, s_t)$
17: **end while**
18: Obtain Pareto front of the posterior mean at the target fidelity $s_\diamond$ subject to $\mathbb{E}[h_q(\boldsymbol{x}, s_\diamond) \mid \mathcal{D}] \geq 0\ \forall q \in [Q]$

for $T \geq \max\{2N_0, 2N_\mathrm{e}\}$. At each t_i, the algorithm queries $\boldsymbol{x}_{t_i} = \arg\max_{\boldsymbol{x}} \sum_j \sigma^2_{t_i-1,j}(\boldsymbol{x}, s_\diamond)$ at the target fidelity.

Step 1 (monotonicity and the greedy choice). For fixed kernel hyperparameters, conditioning on additional observations cannot increase the MF-CGP posterior variance. The configuration $\boldsymbol{x}_{t_i}$ maximizes the sum of posterior variances at the target fidelity before intervention t_i. This sum bounds the corresponding variance sum at any configuration after T interventions. For every $\boldsymbol{x} \in \mathbb{X}$ and $i \leq n$,

$$\sum_j \sigma^2_{T,j}(\boldsymbol{x}, s_\diamond) \;\leq\; \sum_j \sigma^2_{t_i-1,j}(\boldsymbol{x}, s_\diamond) \;\leq\; \sum_j \sigma^2_{t_i-1,j}(\boldsymbol{x}_{t_i}, s_\diamond),$$

Averaging over i,

$$\sup_{\boldsymbol{x}} \sum_j \sigma^2_{T,j}(\boldsymbol{x}, s_\diamond) \;\leq\; \frac{1}{n}\sum_{i=1}^{n}\sum_j \sigma^2_{t_i-1,j}(\boldsymbol{x}_{t_i}, s_\diamond). \tag{7}$$

Step 2 (information-gain bound). Index the initialization and subsequent queries consecutively as $(\boldsymbol{x}_1, s_1), \ldots, (\boldsymbol{x}_T, s_T)$. Let $\boldsymbol{\Sigma}_T$ denote their Gram matrix. For query t, let $\boldsymbol{K}^{(t)}$ denote the $J \times J$ posterior covariance matrix of the outputs conditioned on the first $t-1$ observations. The definition of maximum information gain gives

$$\gamma_T \;\geq\; \frac{1}{2}\ln\det\big(\boldsymbol{I}+\sigma^{-2}\boldsymbol{\Sigma}_T\big) = \frac{1}{2}\sum_{t=1}^{T}\ln\det\big(\boldsymbol{I}+\sigma^{-2}\boldsymbol{K}^{(t)}\big) \;\geq\; \frac{1}{2}\sum_{t=1}^{T}\ln\Big(1+\sigma^{-2}\textstyle\sum_j \sigma^2_{t-1,j}(\boldsymbol{x}_t, s_t)\Big),$$

where the equality follows from the determinant chain rule and the final inequality follows from $\det(\boldsymbol{I} + \boldsymbol{A}) \geq 1 + \mathrm{Tr}(\boldsymbol{A})$ for $\boldsymbol{A} \succeq 0$.

The kernel bound $\tilde{k} \leq 1$ implies $0 \leq \sigma^{-2}\sum_j \sigma^2_{t-1,j}(\boldsymbol{x}_t, s_t) \leq J\sigma^{-2}$. Concavity of the logarithm gives $\ln(1+a) \geq a\,\ln(1+\kappa)/\kappa$ for $a \in [0, \kappa]$, where $\kappa = J\sigma^{-2}$. Applying this inequality with $a = \sigma^{-2}\sum_j \sigma^2_{t-1,j}(\boldsymbol{x}_t, s_t)$ to the information-gain bound above yields

$$\sum_{t=1}^{T}\sum_j \sigma^2_{t-1,j}(\boldsymbol{x}_t, s_t) \;\leq\; \frac{2J}{\ln\big(1+J\sigma^{-2}\big)}\,\gamma_T \;=\; K\,\gamma_T.$$

Step 3 (combine). The sum over $t_1, \dots, t_n$ in Eq. (7) is bounded by the sum over all T queries in Step 2 because posterior variances are nonnegative. Substituting this bound into Eq. (7) and applying $n \geq T/(4N_{\mathrm{e}})$ gives

$$\sup_{\boldsymbol{x}} \|\boldsymbol{\sigma}_T(\boldsymbol{x}, s_\diamond)\|_2^2 \;\leq\; \frac{K\,\gamma_T}{n} \;\leq\; \frac{4\,K\,N_{\mathrm{e}}\,\gamma_T}{T}. \qquad \square$$

Remark B.4. The variance bound holds for any queries selected by VoCI-HVKG between the scheduled target-fidelity queries. Additional observations can only reduce the posterior variance. Without periodic exploration, a misspecified prior may cause a value-directed acquisition to leave some regions unqueried, preventing $\bar{\sigma}_T$ from converging to zero. Theorem 5.8 bounds regret in terms of the realized uncertainty $\bar{\sigma}_T$ without asserting its decay. Lemma B.3 establishes this decay under periodic exploration. The exploration cycle contributes a factor $\sqrt{N_{\mathrm{e}}}$ to the resulting regret bound.

Theorem B.5 (Convergence under exploration)**.** *Let Assumptions 5.1–5.5 hold and fix the kernel hyperparameters. For any $\rho \in (0,1)$, with probability at least $1-\rho$, for all $T \geq \max\{2N_0, 2N_{\mathrm{e}}\}$, Algorithm 3 satisfies*

$$\mathrm{HV}^* - \mathrm{HV}\big[\{\boldsymbol{f}(\boldsymbol{x}, s_\diamond)\}_{\boldsymbol{x}\in\widehat{\mathbb{P}}_T}\big] \;\leq\; 2\big(L+L'\big)\left(\beta_T^{1/2}\sqrt{\frac{4KN_{\mathrm{e}}\gamma_T}{T}} \;+\; \xi\Big(1+\tfrac{2}{\sigma}\sqrt{J\,K\,N_{\mathrm{e}}\,\gamma_T}\Big)\right), \tag{8}$$

and every recommended configuration $\boldsymbol{x} \in \widehat{\mathbb{P}}_T$ satisfies $h_q(\boldsymbol{x}, s_\diamond) \geq -\xi\big(1+\frac{2}{\sigma}\sqrt{JKN_{\mathrm{e}}\gamma_T}\big)$ for all q.

Proof. By Theorem 5.8, with probability at least $1-\rho$, the bound in Eq. (5) holds for all T. Lemma B.3 further gives, for $T \geq \max\{2N_0, 2N_{\mathrm{e}}\}$,

$$\bar{\sigma}_T \;=\; \sup_{\boldsymbol{x}\in\mathbb{X}} \|\boldsymbol{\sigma}_T(\boldsymbol{x}, s_\diamond)\|_2 \;\leq\; \sqrt{\frac{4\,K\,N_{\mathrm{e}}\,\gamma_T}{T}}.$$

The bound on $\bar{\sigma}_T$ gives

$$\beta_T^{1/2}\bar{\sigma}_T \leq \beta_T^{1/2}\sqrt{\frac{4KN_{\mathrm{e}}\gamma_T}{T}},$$
$$\frac{\sqrt{JT}}{\sigma}\bar{\sigma}_T \leq \frac{\sqrt{JT}}{\sigma}\sqrt{\frac{4KN_{\mathrm{e}}\gamma_T}{T}} = \frac{2}{\sigma}\sqrt{JKN_{\mathrm{e}}\gamma_T}.$$

Substituting these two inequalities into Eq. (5) gives Eq. (8). Finally, Theorem 5.8 bounds the constraint violation of every recommended configuration by $\xi\big(1+\frac{1}{\sigma}\sqrt{JT}\,\bar{\sigma}_T\big)$. Applying the same bound on $\bar{\sigma}_T$ gives $\xi\big(1+\frac{2}{\sigma}\sqrt{JKN_{\mathrm{e}}\gamma_T}\big)$, which completes the proof. $\square$

Corollary B.6 (Convergence rate under a well-specified prior)**.** *If $\xi = 0$ and $\gamma_T = o(\sqrt{T})$, all recommended configurations are feasible and*

$$\mathrm{HV}^* - \mathrm{HV}\big[\{\boldsymbol{f}(\boldsymbol{x}, s_\diamond)\}_{\boldsymbol{x}\in\widehat{\mathbb{P}}_T}\big] = \mathcal{O}\Big(MR^{M-1}\big(B+\sqrt{\gamma_T}\big)\sqrt{N_{\mathrm{e}}\,\gamma_T/T}\Big) \;\xrightarrow[T\to\infty]{}\; 0.$$

For the RBF kernel over the $d{+}1$ input dimensions, $\gamma_T = \mathcal{O}\big((\ln T)^{d+2}\big)$, giving a rate of $\tilde{\mathcal{O}}(1/\sqrt{T})$.

Proof. Setting $\xi = 0$ in Theorem B.5 removes the misspecification term and makes the constraint-violation bound zero. Thus, all recommended configurations are feasible. The hypervolume bound becomes $\mathrm{HV}^* - \mathrm{HV}_{\boldsymbol{f}}[\widehat{\mathbb{P}}_T] \leq 2(L+L')\beta_T^{1/2}\sqrt{4KN_{\mathrm{e}}\gamma_T/T}$ with $\beta_T^{1/2} = B + \sqrt{2(\gamma_T + 1 + \ln(1/\rho))}$. Substituting $L = 2MR^{M-1}$ and keeping the remaining problem-dependent constants fixed gives the stated order. Since $\gamma_T = o(\sqrt{T})$, both $\sqrt{\gamma_T/T}$ and $\gamma_T/\sqrt{T}$ converge to zero. Thus the hypervolume regret also converges to zero.

For the RBF kernel, $k_{\mathrm{in}}k_{\mathrm{fid}}$ is an RBF kernel over the $d{+}1$ continuous input dimensions. Its information gain satisfies $\gamma_T = \mathcal{O}((\ln T)^{d+2})$ on a compact domain (Srinivas et al., 2010). The causal component $\langle\widehat{\boldsymbol{\sigma}}, \widehat{\boldsymbol{\sigma}}'\rangle$ is a finite-rank kernel and contributes $\mathcal{O}(\ln T)$, while the index kernel over the J outputs contributes a constant factor. The compositional bounds for sums and products of kernels preserve the polylogarithmic information gain, giving the rate $\tilde{\mathcal{O}}(1/\sqrt{T})$ and completing the proof. $\square$

Remark B.7. Under Algorithm 3, the misspecification term of Theorem 5.8 is of order $\xi\sqrt{\gamma_T}$. For the RBF kernel, this term grows only polylogarithmically in T and does not accumulate linearly. Corollary 5.9(ii) also allows Algorithm 3 to use the target-slice posterior. If each exploration query selects the point of maximum posterior uncertainty under this posterior, the algorithm performs $n = \lfloor (T - N_0)/N_{\mathrm{e}} \rfloor$ target-fidelity interventions. The greedy choice of Lemma B.3 then holds for the restricted posterior. Therefore, Theorem B.5 applies to the GP over $\mathbb{X} \times \{s_\diamond\}$, with its restricted kernel and information gain and with T replaced by n. This guarantee does not require any assumption on the auxiliary fidelities.

B.7 Discussion

Role of the causal prior Lemma 5.7 quantifies the benefit of the causal prior in two ways. First, the RKHS bound B applies to the residual $\bar{\boldsymbol{f}} - \bar{\boldsymbol{f}}_{\mathrm{do}} - \boldsymbol{\Delta}$ rather than to $\bar{\boldsymbol{f}}$ itself. Thus, an informative CPM reduces the RKHS norm that appears in the confidence bound. Second, when $\xi > 0$, the misspecification term scales with ξ, which bounds the uniform error of the causal estimates. Improving the observational data or the fit of the causal mechanisms directly reduces this term. Since the CPM models the fidelity dependence of the outputs, B and ξ can remain small even when auxiliary sources are biased.

Role of the exploration cycle Theorem 5.8 holds for every query sequence, but its bound converges to zero only if $\bar{\sigma}_T \to 0$. A value-directed acquisition alone does not guarantee this condition. If the posterior assigns low uncertainty to an inaccurate region, the acquisition may not query that region and the prediction error can persist. Algorithm 3 periodically queries the point of maximum uncertainty at the target fidelity and guarantees uniform posterior contraction. This gives the rate in Theorem B.5, with a factor $\sqrt{N_{\mathrm{e}}}$ in the bound.

Interpretation of Corollary 5.9 The corollary does not state that every low-fidelity evaluation is useful. It states that the guarantee does not depend on the worst-case cross-fidelity bias. The causal prior models the cross-fidelity structure, so the MF-CGP does not have to infer this structure only by extrapolating from low-fidelity observations. VoCI-HVKG selects each configuration–fidelity pair by its cost-normalized expected improvement in the feasible hypervolume at the target fidelity. Theorem 5.8 quantifies how causal-prior misspecification affects the regret bound through ξ.

Relation to rMFBO Mikkola et al. (2023) bound the regret difference between their robust MFBO method and SFBO by $\varepsilon \max\{T\widehat{M}_T d^{T+1}, 2\}$. This bound is exponential in T through the input dimension d, and the authors note that it is useful only in the first rounds. They do not provide a convergence rate. In comparison, Theorem 5.8 gives a guarantee at every iteration, with the dependence on auxiliary fidelities confined to the constants of Assumption 5.2 (Corollary 5.9). For Algorithm 3, Theorem B.5 gives a sublinear convergence rate when $\xi = 0$ (Corollary B.6) and an additive $\xi\sqrt{\gamma_T}$ misspecification term otherwise.