Confidence-Based Handover for Causal Bayesian Optimization with Adaptive Expert Trust

Abstract

Causal Bayesian Optimization (CBO) has emerged as a powerful paradigm for decision-making under uncertainty, effectively leveraging causal structures to guide interventions. However, existing CBO methods do not adequately integrate domain-specific expertise, potentially leading to inefficient exploration and suboptimal outcomes in real-world applications involving complex human decision-making and economic considerations. To bridge this gap, we propose Expert-guided Causal Bayesian Optimization (ECBO), a framework that integrates human expert knowledge into the CBO process. We introduce an expert weighting mechanism that adaptively modulates the Expected Improvement (EI) acquisition function, reflecting expert endorsements or exclusions of candidate solutions. By embedding these expert-derived preferences, our model proactively balances exploration and exploitation. Additionally, we implement a robust harm-free mechanism to guard against potentially detrimental expert interventions, alongside an adaptive trust parameter that dynamically adjusts expert influence based on observed discrepancies. Our experimental results underscore the practical value of effectively incorporating expert knowledge in complex causal environments.

1 Introduction

Bayesian Optimization (BO) is widely used for optimizing expensive black-box functions. The concept originated from the works of [26, 27, 28], but it gained broader attention following its popularization by [1]. In traditional BO, the surrogate model heavily relies on data-driven covariance structures, defined via its kernel, to assess variable relevance and interactions. This approach, however, does not account for external prior knowledge concerning the roles or semantics of variables, effectively ignoring causal structures. Additionally, standard BO intervenes on all variables simultaneously, disrupting the natural dependencies of the system and potentially leading to suboptimal or infeasible interventions.

To address these issues, [2] introduced Causal Bayesian Optimization (CBO), which integrates causal knowledge, often represented as a Structural Causal Model (SCM) [3], into the BO framework. In this setup, "do-interventions" are applied to select variables, propagating through the causal graph, while non-intervened variables follow observational distributions. By embedding the graph structure into the Gaussian Process prior and constraining the acquisition to valid intervention subspaces, CBO reduces the search space, enhancing global decision-making in causal contexts, and avoiding suboptimal solutions that might arise from standard BO when causal propagation is neglected. Recently, various extensions have emerged. Notable contributions include [13, 4, 19, 12, 30, 31, 29, 32]. A summary of these advancements is provided in the Appendix 2.

Despite these advances, existing CBO frameworks often underuse a critical resource: human domain expertise [7, 8, 9]. Practitioners in chemical engineering, clinical medicine, and robotics [24, 25] possess rich mental models that encode physical constraints and heuristic rules (regulatory caps, graduated dosing, stability bounds). **Why expert input?** It supplies priors over where to look and what to avoid—information that the SCM does not encode, such as feasibility, risk, and utility trade-offs. **How does it help CBO?** It reweights candidates, improving early sample efficiency and reducing avoidable regret and steering the optimizer away from well-known dead ends and toward promising regimes faster. However, if incorporated naively, imprecise or inconsistent advice can mislead the surrogate or violate causal assumptions. The open challenge, therefore, is to integrate expert guidance effectively without compromising CBO's theoretical guarantees.

Recent human-AI collaborative works [6, 7, 8, 9, 23] have demonstrated that embedding human feedback into a BO framework can accelerate search. However,

some fundamental limitations remain. First, expert input is often reduced to binary or global scores, which cannot express variable- and region-specific preferences needed to balance, e.g., catalyst load, solvent volatility, and safety limits. Second, human feedback is noisy and nonstationary. Treating it as a single fixed-weight reward miscalibrates the surrogate and triggers policy oscillations and long-tail regret. In addition, generic feedback may propose causally infeasible actions and corrupt learning, excluding valid interventions within the context of Causal Bayesian Optimization (CBO), thereby violating identifiability assumptions and potentially leading to biased or slower model updates. These limitations underscore the need for a framework that supports fine-grained expert preferences to guide intervention strategies.

To address these fundamental limitations, we propose the ECBO framework, which systematically integrates dimension-specific, real-valued expert feedback into the causal optimization process. Our approach assigns nuanced expert preference weights to each decision variable individually, capturing a rich spectrum of domain knowledge. Practitioners can articulate preferences such as gradually promoting a particular variable, moderately suppressing another, or strictly excluding unsafe regions, effectively translating expert insights into explicit optimization constraints. ECBO searches over executable intervention subspaces and then lifts the chosen setting to a full decision, so every proposal is implementable.

Building on this expert-guided foundation, ECBO employs an expert-enhanced acquisition that blends classic Expected Improvement (EI) with expert preference weights via an adaptive trust coefficient. This trust coefficient is updated based on a decision-aligned improvement signal; it rises when expert guidance helps and falls when it hurts. A conservative safety gate filters low-confidence expert regions, and a confidence-based handover automatically disables the expert once its surrogate becomes reliable. The scheme is soft-weighted, circumventing rigid binary rejection by gently steering exploration, so that informative candidates are preserved and late-stage oversteer is avoided. In addition, we introduce the average GAP metric to quantitatively evaluate the trajectory-level efficiency of our approach compared to existing baselines. In short, the main contributions of this paper are summarized as follows:

- We propose an adaptive, expert-weighted EI acquisition function that dynamically blends traditional Expected Improvement with expert preference weights.
- We introduce a trust calibration mechanism that updates the expert trust coefficient online, growing it when expert advice improves the objective and shrinking it when advice proves harmful.
- We introduce PA-GAP, a path-aware efficiency metric that multiplies normalized improvement by a time weight and averages over rounds.
- We prove no-harm relative to a coupled EI baseline and same-order convergence/regret under standard GP assumptions, with guarantees that match the algorithm actually used.

2 Background and Problem Setting

In this section, we first formalize the interventional optimization problem over a structural causal model, then show how to prune the combinatorial intervention-set space to a tractable family, and finally introduce the causal GP surrogate that underpins all downstream acquisition strategies.

2.1 Problem Formulation and Intervention-Set Construction

We consider a known structural causal model (SCM) [3] $\mathcal{M} = \langle \mathcal{G}, \mathbf{U}, \mathbf{V}, \mathcal{F}, P(\mathbf{U}) \rangle$, where $\mathcal{G}$ is a directed acyclic graph (DAG), $\mathbf{U}$ is the set of exogenous (noise) variables, $P(\mathbf{U})$ its exogenous distribution, $\mathbf{V}$ is the set of endogenous variables, and $\mathcal{F}$ is the collection of structural equations $V_i = f_i(\mathrm{pa}_i, U_i)$. Within the set of endogenous variables $\mathbf{V}$, we denote a subset of controllable variables $X = (X_1, \ldots, X_d) \subset \mathbf{V}$, and a target outcome variable $Y \in \mathbf{V}$.

An (possibly partial) intervention on a subset of variables $S \subseteq X$ involves enforcing the assignment $S = \mathbf{x}_s$, where $\mathbf{x}_s$ is a value from the corresponding domain $\mathcal{X}_S$, which is the Cartesian product of the domains of X_i in S. This action results in a post-interventional distribution for the outcome variable, $P(Y \mid \mathrm{do}(S = \mathbf{x}_s))$. The objective is to identify an optimal intervention $(S^\star, \mathbf{x}_s^\star)$ that maximizes the expected value of the outcome Y, subject to a finite budget T. This optimization problem is formulated as:

$$(S^\star, \mathbf{x}_s^\star) = \arg \max_{S \subseteq X,\, \mathbf{x}_s \in \mathcal{X}_S} \mu_Y(S, \mathbf{x}_s),$$

$$\text{where} \quad \mu_Y(S, \mathbf{x}_s) := \mathbb{E}[\, Y \mid \mathrm{do}(S = \mathbf{x}_s)\,] \qquad (1)$$

In practice, the intervention space is exponential ($2^{|X|}$), so we limit our search to a more manageable, predefined *Intervention Search Space*, denoted by $\mathcal{S}_{\text{exp}} \subseteq \mathcal{P}(X)$.

Define a subset $S \subseteq X$ as a *Potentially Optimal Set (POS)* if it satisfies both:

Redundancy-Free: There is no strict subset $S' \subset S$ and assignment $s' \in \mathcal{X}_{S'}$ such that

$$\max_{s \in \mathcal{X}_S} \mu_Y(S, s) = \max_{s' \in \mathcal{X}_{S'}} \mu_Y(S', s'). \tag{2}$$

That is, no variable in S can be removed without potentially reducing the best achievable outcome.
Potential Optimality: There exists a realization of the structural functions and exogenous noise such that

$$\max_{s \in \mathcal{X}_S} \mu_Y(S, s) > \max_{w \in \mathcal{X}_W} \mu_Y(W, w) \quad \text{for all } W \in \mathcal{S}_{\text{rf}} \setminus \{S\}. \tag{3}$$

where $\mathcal{S}_{\text{rf}}$ denotes the family of all redundancy-free sets. We consider the finite family $\mathcal{S}_{\text{exp}}$ of all potential optimal sets as the *Intervention Search Space*.

2.2 Surrogate Modeling under Causal Structure

We model the interventional response surface using a Gaussian Process (GP). For each intervention set $S \in \mathcal{S}_{\text{exp}}$, we maintain an independent GP $\mathcal{G}_S$ mapping $\mathbf{x}_S \in \mathcal{X}_S$ to scalar outcome Y.

Gaussian Processes in brief. A Gaussian Process (GP) prior over a scalar function $f(\cdot)$ is specified by mean $m(\cdot)$ and kernel $k(\cdot, \cdot)$; given inputs X and outputs $\mathbf{y}$, the posterior mean and variance at $\mathbf{x}$ follow the standard closed forms. For each intervention set $S \in \mathcal{S}_{\text{exp}}$, we model the interventional response

$$f_S(\mathbf{x}_S) = \mathbb{E}[Y \mid \text{do}(S = \mathbf{x}_S)], \tag{4}$$

with an independent GP $\mathcal{G}_S$ defined on $\mathcal{X}_S$:

$$f_S(\cdot) \sim \mathcal{GP}\big(m_S(\cdot),\, k_S(\cdot, \cdot)\big). \tag{5}$$

Let the dataset at round t be $\mathcal{D}_t^S = \{(\mathbf{x}_S^{(i)}, y^{(i)})\}_{i=1}^{n_t^S}$. Let $X_S = [\mathbf{x}_S^{(i)}]$, $\mathbf{y}_S = [y^{(i)}]$, and $K_S = k_S(X_S, X_S)$, then the posterior of $\mathcal{G}_S$ at $\mathbf{x}_S \in \mathcal{X}_S$ is

$$\mu_t^S(\mathbf{x}_S) = m_S(\mathbf{x}_S) + k_S(\mathbf{x}_S, X_S)\big(K_S + \sigma^2 I\big)^{-1}\big(\mathbf{y}_S - m_S(X_S)\big) \tag{6}$$

$$\sigma_t^{2,S}(\mathbf{x}_S) = k_S(\mathbf{x}_S, \mathbf{x}_S) - k_S(\mathbf{x}_S, X_S)\big(K_S + \sigma^2 I\big)^{-1} k_S(X_S, \mathbf{x}_S) \tag{7}$$

Gradients. Assume k_S and m_S are differentiable. Define

$$\alpha_S = (K_S + \sigma^2 I)^{-1}\big(\mathbf{y}_S - m_S(X_S)\big), \tag{8}$$

$$C_S = (K_S + \sigma^2 I)^{-1}, \tag{9}$$

Then, we obtain

$$\nabla_{\mathbf{x}_S} \mu_t^S(\mathbf{x}_S) = \nabla_{\mathbf{x}_S} m_S(\mathbf{x}_S) \; + \; \nabla_{\mathbf{x}_S} k_S(\mathbf{x}_S, X_S)\, \alpha_S, \tag{10}$$

$$\nabla_{\mathbf{x}_S} \sigma_t^{2,S}(\mathbf{x}_S) = \nabla_{\mathbf{x}_S} k_S(\mathbf{x}_S, \mathbf{x}_S) - 2\, \nabla_{\mathbf{x}_S} k_S(\mathbf{x}_S, X_S)\, C_S\, k_S(X_S, \mathbf{x}_S) \tag{11}$$

For stationary kernels $k_S(\mathbf{x}_S, \mathbf{x}_S)$ that are constant, the first term in (Eq.11) is zero. These derivatives are used in Section 3 to optimize the acquisition.

3 Methodology

3.1 Expert Preference Representation

We incorporate human or domain expertise through a lightweight query protocol that returns structured preference signals for any candidate intervention $\mathbf{X} = \mathbf{x}$. Each query yields one of four outcomes:(i) **Exclude,** down-weight the coordinate and any candidate that contains it; (ii) **Promote,** prefer increasing the coordinate; (iii) **Suppress,** prefer decreasing the coordinate; (iv) **Uncertain,** no directional guidance.

Preference Modeling. We convert these heterogeneous signals into a smooth preference weight $w(\mathbf{x}) \in [0, \infty)$ that modulates the acquisition function (Section 3.2). Let $\mathcal{K}$ denote the set of expert-specified rules. For each rule $k \in \mathcal{K}$, the expert provides: (i)a tagging function τ_k, mapping intervention feature to discrete tags (ii) and a tag-to-weight map ω_k, assigning a multiplicative scalar to each tag.

The final weight $w(\mathbf{x})$ is computed by aggregating the per-rule contributions multiplicative:

$$w(\mathbf{x}) = \prod_{\substack{k \in \mathcal{K} \\ x_j \in \mathbf{x}}} \omega_k\left(\tau_k(x_j)\right), \tag{12}$$

a concrete example is presented in the Appendix 3.

Uncertainty Handling Rule. A critical design is how to react when all feedback within a given intervention subset S is marked "uncertainty" during an iteration. To prevent spurious shifts driven by noise-dominated gradients, we skip the expert update in that subset: neither the acceptance model p_θ (defined in Section 3.3) nor the trust coefficient η_t (defined in Section 3.2) is revised. This rule preserves stability, especially on datasets where uncertain signals are frequent, and empirically prevents degradation observed when forcing updates with empty supervision.

3.2 Expert-Weighted Causal Expected Improvement

We build upon the standard EI criterion to incorporate expert guidance. Let $y^+ = \max\{y^{(i)}\}$ be the best

observed outcome so far and denote the GP posterior mean and standard deviation at $\mathbf{x}$ by $\mu(\mathbf{x})$ and $\sigma(\mathbf{x})$. Let $f(\mathbf{x}) = \mathbb{E}[Y \mid do(\mathbf{X} = \mathbf{x})]$. The classic EI for maximization is defined as

$$\mathrm{EI}(\mathbf{x}) \;=\; \mathbb{E}\big[\max\{0,\, f(\mathbf{x}){-}y^{+}\}\big] \;=\; \sigma(\mathbf{x})\big(z, \Phi(z){+}\phi(z)\big),$$
$$\text{where} \quad z = \frac{\mu(\mathbf{x}) - y^{+}}{\sigma(\mathbf{x})}, \tag{13}$$

where Φ and ϕ are the standard normal CDF and PDF, respectively.

To inject expert preference and intervention expense, we modulate EI by the trust-weighted deviation of the expert weight $w(\mathbf{x})$ from unity. We define the Expert-weighted Causal EI (ECEI) as

$$\mathrm{ECEI}(\mathbf{x}) = \frac{\big[\,1 + \eta_t\,(w(\mathbf{x}) - 1)\big]\,\mathrm{EI}(\mathbf{x})}{co(\mathbf{x})}, \tag{14}$$

where $\eta_t \in [0, 1]$ is the trust coefficient at iteration t and $co(\mathbf{x}) > 0$ denotes the cost of executing intervention $\mathbf{X} = \mathbf{x}$, capturing resources or time required.

Its gradient, which drives our gradient-based acquisition optimizer, decomposes naturally into two terms:

$$\begin{aligned}\nabla\mathrm{ECEI}(\mathbf{x}) = &\frac{\big[\,1 + \eta_t\,(w(\mathbf{x}) - 1)\big]\,\nabla\mathrm{EI}(\mathbf{x})}{co(\mathbf{x})} \\ &+ \frac{\eta_t\,\mathrm{EI}(\mathbf{x})\,\nabla w(\mathbf{x})}{co(\mathbf{x})} \\ &- \frac{\big[\,1 + \eta_t\,(w(\mathbf{x}) - 1)\big]\,\mathrm{EI}(\mathbf{x})}{co(\mathbf{x})^2}\,\nabla co(\mathbf{x}).\end{aligned} \tag{15}$$

The first term of the numerator rescales the standard EI gradient, while the second injects directional signals from the expert model. Finally, each intervention subset $S \in \mathcal{S}_{\exp}$ maintains its own surrogate and expert-weighted acquisition. During each BO iteration, we optimize $\mathrm{ECEI}_S(\mathbf{x}_S)$ separately in parallel and then compare the best proposals across all $\mathcal{S}_{\exp}$ to select the next evaluation point.

3.3 Trust Modulation and No-harm Weighting

We introduce a scalar trust coefficient $\eta_t \in [0, 1]$ to balance surrogate-driven search and expert guidance. Intuitively, we wish to increase trust when expert advice yields better candidate points, and decrease it when advice harms expected improvement.

Rationale for EI Difference To quantify "potential harm" or benefit of expert guidance, we compare the classic EI with our ECEI. Define the batch-average difference

$$\overline{\Delta\mathrm{EI}}_t \;=\; \frac{1}{|S|}\sum_{\mathbf{x}\in S}\Big[\;\mathrm{ECEI}_t(\mathbf{x}) - \mathrm{EI}_t(\mathbf{x})\;\Big]. \tag{16}$$

For the selected candidate $\mathbf{x}_t$, we use the single-point surrogate $\Delta EI_t := EI_t(\mathbf{x}_t)\,[\,w_t(\mathbf{x}_t) - 1\,]$. If $\overline{\Delta\mathrm{EI}}_t < 0$, expert weighting reduces EI (a sign of harm or misalignment), whereas $\overline{\Delta\mathrm{EI}}_t > 0$ indicates a net improvement from expert input. ΔEI is the right signal because it measures the expert's marginal effect on the exact quantity we optimize. Holding the posterior and candidate set fixed, it isolates how expert weighting changes decisions.

Trust Update We update trust proportionally to this measured effect:

$$\eta_{t+1} = \;\mathrm{clip}\Big(\eta_t \;+\; \gamma\,\overline{\Delta\mathrm{EI}}_t,\; 0, 1\Big), \tag{17}$$

where $\gamma > 0$ is a learning rate and $\mathrm{clip}(\eta, 0, 1) = \min\{\max\{\eta, 0\}, 1\}$. Thus, beneficial advice ($\overline{\Delta\mathrm{EI}}_t > 0$) will increase the update term (increasing net trust), while harmful advice ($\overline{\Delta\mathrm{EI}}_t < 0$) will reduce trust.

Safe-Check We adopt a calibrated reliability gate based on the expert-confidence weight $w(\mathbf{x})$:

$$\mathrm{Safe}(\mathbf{x}) \;=\; \mathbf{1}_{\{\,w(\mathbf{x}) \geq \alpha\}}, \tag{18}$$

When $\mathrm{Safe}(\mathbf{x}_t) = 0$ the expert is disabled for this round and we set $\eta_{t+1} = 0$ (fallback to the baseline EI), high-confidence regions and stabilizes optimization.

Handover-check We implement the expert surrogate as a simple logistic classifier

$$p_\theta(\text{accept} \mid \mathbf{x}) = \sigma\big(\phi(\mathbf{x})^\top \beta\big), \tag{19}$$

trained on the labeled expert queries $\mathcal{D}_t = \{(\mathbf{x}^{(i)}, r^{(i)})\}_{i=1}^{n_t}$ with $r^{(i)} \in \{0, 1\}$. Let uncertainty$(\mathbf{x})$ denote the surrogate's predictive uncertainty (returned by the implementation).

We define a global uncertainty upper bound at round t as

$$U_t \;=\; \max_{(\mathbf{x}, r)\in\mathcal{D}_t} \text{uncertainty}(\mathbf{x}). \tag{20}$$

If $U_t < g_{\mathrm{thr}}$ we stop using the expert signal and set $\eta_{t+1} = 0$, otherwise the expert remains active.

Harm-Free Principle Our design aims to be harm-free with respect to a coupled EI baseline: when expert guidance helps, it is used; when it hurts, its influence is disabled. Concretely, we combine (i) a decision-aligned trust update driven by ΔEI (Eq.17), (ii) a conservative gate that disables expert input in low-confidence regions (Eq.18), and (iii) a confidence-based handover that turns the expert off once its surrogate becomes reliable (Eq.20). This is a performance guarantee relative to EI under bounded weights and costs and a finite candidate set. Formal proofs appear in Section 3.5.

Algorithm 1 Expert Causal Bayesian Optimization

Input: observational data $\mathcal{D}_0^O$, interventional data $\mathcal{D}_0^I$, subset family $\mathcal{S}_{\text{exp}}$
Parameter: budget T
Output: optimal subset $S^\star$, assignment $\mathbf{x}^\star$, estimate $\widehat{\mathbb{E}}[Y^\star \mid \text{do}(S^\star = \mathbf{x}^\star)]$

for $t = 1$ **to** T **do**
  $\mathcal{C}_t \leftarrow \emptyset$ # candidate set
  for each $S \in \mathcal{S}_{\text{exp}}$ **do**
    $\mathbf{x}_S^\star \leftarrow \arg\max_{\mathbf{x}_S \in \mathcal{X}_S} \text{ECEI}_t^S(\mathbf{x}_S)$
    $v_S \leftarrow \text{ECEI}_t^S(\mathbf{x}_S^\star)$
    $\mathcal{C}_t \leftarrow \mathcal{C}_t \cup \{(S, \mathbf{x}_S^\star, v_S)\}$
  end for
  $(S_t, \mathbf{x}_t, v_t) \leftarrow \arg\max_{(S,\mathbf{x},v)\in\mathcal{C}_t} v$
  $y_t \leftarrow \mathbb{E}[Y \mid do(S_t = \mathbf{x}_t)]$
  if not all coordinates in S_t are *uncertain* **then**
    $\Delta\text{EI}_t \leftarrow \text{EI}_t(\mathbf{x}_t)\,[\,w_t(\mathbf{x}_t) - 1\,]$
    $\eta_{t+1} \leftarrow \text{clip}\big(\eta_t + \gamma\,\overline{\Delta\text{EI}}_t,\, 0, 1\big)$ *(trust update)*
    if $w_t(\mathbf{x}_t) < \alpha$ **then**
      $\eta_{t+1} \leftarrow 0$
    end if *(safe-check)*
    if $U_t < g_{\text{thr}}$ **then**
      $\eta_{t+1} \leftarrow 0$
    end if *(handover)*
    label $\leftarrow$ query_human$(\mathbf{x}_t)$
    update expert surrogate with $(\mathbf{x}_t, \text{label})$
  end if
  update **only** $\mathcal{G}_{S_t}$ with $(\mathbf{x}_t, y_t)$
end for
return $S^\star, \mathbf{x}^\star, \widehat{\mathbb{E}}[Y^\star \mid \text{do}(S^\star = \mathbf{x}^\star)]$

3.4 Iterative Loop and Post-update Policy

Our algorithm proceeds in rounds $t = 1, \ldots, T$. At each iteration, we select a candidate subset $S_t \in \mathcal{S}_{\text{exp}}$ and optimize the ECEI to propose an intervention $\mathbf{x}_t$. We then evaluate $y_t = \mathbb{E}[Y \mid do(\mathbf{X} = \mathbf{x}_t)]$ and solicit expert feedback on $\mathbf{x}_t$. If in the previous Expert Eclicitation, the expert expresses fully neutral preferences (i.e., uncertainty across all dimensions in the current subset), we skip the preference model update. Otherwise, we update the preference model accordingly: we adjust the trust coefficient η_t based on the average deviation in expert weights, perform safe-check and handover-check. If both check passed, we use intervention $\mathbf{x}_t$ and the queried label to update the expert logistic surrogate classifier. Finally, we update the GP posterior using the new observation $(\mathbf{x}_t, y_t)$. This iterative loop ensures that surrogate modeling, expert guidance, and trust modulation work in concert to enable efficient and causally grounded exploration.

3.5 Theoretical Guarantees

Assumptions (A1) The objective is linearly normalized to $[0, 1]$.

(A2) Costs and expert weights are bounded:

$$0 < c_{\min} \le \text{co}(\mathbf{x}) \le c_{\max} < \infty,$$
$$0 < w_{\min} \le w(\mathbf{x}) \le w_{\max} < \infty. \tag{21}$$

(A3) Safety- and handover-check do not increase η_t (we do not require η_t to be monotone).

(A4) At each round, the implementation optimizes the acquisition over a finite candidate pool $\mathcal{X}_t^{\text{cand}} \subseteq \mathcal{X}$; the solver returns a ρ_t-approximate maximizer, i.e.,

$$\text{ECEI}_t(\mathbf{x}_t^{\text{ec}}) \;\ge\; \rho_t \max_{\mathbf{x}\in\mathcal{X}_t^{\text{cand}}} \text{ECEI}_t(\mathbf{x}), \rho_t \in (0, 1]. \tag{22}$$

Here $\mathbf{x}_t^{\text{ec}} \in \mathcal{X}_t^{\text{cand}}$ is the ECBO proposal at round t.

Coupled baseline Let $\mathcal{B}^{\mathcal{H}}$ be the baseline that runs on the same history and uses EI as the acquisition (with $\eta \equiv 0$):

$$\mathbf{x}_t^{\text{base}} \in \arg\max_{\mathbf{x}\in\mathcal{X}_t^{\text{cand}}} \text{EI}_t(\mathbf{x}). \tag{23}$$

Here $\mathbf{x}_t^{\text{base}} \in \mathcal{X}_t^{\text{cand}}$ is the coupled EI baseline choice on the same history and pool, obtained by maximizing EI_t (equivalently, setting $\eta \equiv 0$ in Eq. (14)). It serves as the comparator for no-harm analysis.

Lemma 1 (Proportional bounds between ECEI and EI). For any $\mathbf{x}$ and t, from Eq. (14),

$$\alpha_{\min,t}\,\text{EI}_t(\mathbf{x}) \;\le\; \text{ECEI}_t(\mathbf{x}) \;\le\; \alpha_{\max,t}\,\text{EI}_t(\mathbf{x}), \tag{24}$$

where

$$\alpha_{\min,t} = \frac{1 + \eta_t\big(w_{\min} - 1\big)}{c_{\max}}, \quad \alpha_{\max,t} = \frac{1 + \eta_t\big(w_{\max} - 1\big)}{c_{\min}} \tag{25}$$

Theorem 1 (Per-round competitiveness vs. baseline). Define

$$\kappa_t \;:=\; \rho_t\,\frac{\alpha_{\min,t}}{\alpha_{\max,t}} \;=\; \rho_t\,\frac{c_{\min}}{c_{\max}}\cdot\frac{1 + \eta_t\big(w_{\min} - 1\big)}{1 + \eta_t\big(w_{\max} - 1\big)} \;\in (0, 1]. \tag{26}$$

Then for every t,

$$\text{EI}_t\big(\mathbf{x}_t^{\text{ec}}\big) \;\ge\; \kappa_t \cdot \max_{\mathbf{x}\in\mathcal{X}_t^{\text{cand}}} \text{EI}_t(\mathbf{x}) \;=\; \kappa_t \cdot \text{EI}_t\big(\mathbf{x}_t^{\text{base}}\big). \tag{27}$$

Corollary 1 (Uniform constant). If $\rho_t \ge \rho_{\min} > 0$, then

$$\kappa_t \;\ge\; \kappa_{\min} \;:=\; \rho_{\min}\,\frac{c_{\min}}{c_{\max}}\,\frac{w_{\min}}{w_{\max}} \;>\; 0. \tag{28}$$

Theorem 2 (Conditional coupling; harm-free). Suppose at some round τ we trigger $\eta_\tau = 0$ (e.g., via safety/handover zeroing or Eq. (17) with clipping). If, in addition, the cost co is fixed (independent of $\mathbf{x}$) thereafter, then for all $t \geq \tau, \mathbf{x}_t^{\text{ec}} = \mathbf{x}_t^{\text{base}}$. Consequently, for any horizon $T \geq \tau$, the subsequent trajectories and losses coincide.

Theorem 3 (Inherited convergence and regret). Assume the coupled baseline $\mathcal{B}^{\mathcal{H}}$ (with the given kernel and noise) satisfies best-so-far simple regret $S_{\mathcal{B}}(T) \to 0$ and/or cumulative regret $\mathcal{R}_{\mathcal{B}}(T) = o(T)$. Then ECEI also satisfies

$$S_{\text{ECBO}}(T) \to 0 \quad \text{and/or} \quad \mathcal{R}_{\text{ECBO}}(T) = o(T). \tag{29}$$

Remarks. (R1) Theorem 1 provides per-round hard lower bounds that depend explicitly on solver accuracy and the boundedness of co and w. Any $w_{\min} > 0$ and $\rho_{\min} > 0$ yield a global constant $\kappa_{\min}$.
(R2) Theorem 2 is a sufficient-condition coupling and does not require η_t to be monotone.
(R3) Theorem 3 gives conditional inheritance (matching the known properties of the EI baseline); we refrain from information-gain-style high-probability claims to avoid a mismatch with the implementation[1].

4 Experiment

4.1 Datasets

We evaluate our method across five scenarios: ToyGraph, CompleteGraph, and PSA (synthetic datasets generated from predefined SCMs [2, 17]) and Diabetes and PSA-CDC (real datasets with SCMs inferred using LiNGAM [14]). The details of each dataset are provided in Appendix 7.

4.2 Baselines

We compare ECBO with several baselines: standard BO [1], CBO [2], COBOL (expert-labeling enhanced BO) [6], constrained CBO (cCBO) [19], dynamic CBO (DCBO) [13] for non-stationary SCM, model-based CBO (MCBO) [4], and high-dimensional CBO (HCBO) [5]. Full details of these baselines are provided in Appendix 6.

4.3 Performance metric

We evaluated all methods with two numerical metrics and visualization metrics: GAP, Path-Aware GAP (PA-GAP), and best-so-far simple regret.

[1]Proof of Section 3.5 is included in the Appendix 4

GAP was introduced in DCBO [13] to quantify optimization efficiency, which was defined as:

$$GAP = \frac{\overbrace{\frac{y(\mathbf{x}_t^*) - y(\mathbf{x}_{\text{init}})}{y^* - y(\mathbf{x}_{\text{init}})}}^{\text{Improvement Term}} + \overbrace{\frac{T - t^*}{T}}^{\text{Efficiency Term}}}{\underbrace{1 + \frac{T-1}{T}}_{\text{Normalization}}}, \tag{30}$$

where $y(\mathbf{x}) = \mathbb{E}[Y \mid do(\mathbf{X} = \mathbf{x})]$ denotes the objective function, y^* is the global optimum, and $\mathbf{x}_{\text{init}}$ and $\mathbf{x}_t^*$ are the initial and best-evaluated points at trial t, respectively. T is the total number of trials, and t^* denotes the trial index at which $y(\mathbf{x}_t^*)$ is discovered. It equals to T if the entire optimization has no improvement[2]. It incorporates an improvement term, capturing the ratio of achieved improvement to the maximum possible, and an efficiency term, rewarding early discovery of the optimum.

However, it pays limited attention to the overall optimization process, as it only uses the final result in the calculation, and it tends to over-penalize later trials, even when they achieve better optimal values. To address these issues, we introduce **PA-GAP** as:

$$\text{PA-GAP} = \frac{1}{T}\sum_{t=1}^{T}\left(\underbrace{\frac{y(\mathbf{x}_t^*) - y(\mathbf{x}_{\text{init}})}{y^* - y(\mathbf{x}_{\text{init}})}}_{\text{Improvement Term}} * \underbrace{\frac{T - (t-1)}{T}}_{\text{Efficiency Term}}\right). \tag{31}$$

By construction, each trial's score lies in the range [0, 1], where 0 indicates no improvement over the initialization and 1 suggests that the optimum is found in the first trial. Larger values correspond to faster and more effective search. Averaging across trials rewards algorithms that perform well throughout the optimization process, rather than only converging early or late in the process.

The **best-so-far simple regret** $\mathcal{R}$ [35] is defined as:

$$\mathcal{R} = y(\mathbf{x}_t^*) - y^*, \tag{32}$$

which describes how close the current optimal value is to the best theoretical value.

4.4 Main Results

Each experiment is repeated twenty times with different random seeds and independent initialization of

[2]DCBO originally sets $t^* = 0$ when no improvement occurs, which restricts GAP to $\left[\frac{T}{2T-1}, 1\right]$ rather than $[0, 1]$. We correct this by setting $t^* = T$. A detailed example is provided in Appendix 8.1.

Table 1: Performance across eight models on five datasets. Each value is average ± std. across twenty random initializations. Best values per dataset are in **bold**.

Metrics	Model	ToyGraph	CompleteGraph	PSA	Diabetes	PSA_CDC
PA-GAP	BO	0.224±.139	0.198±.155	0.125±.064	0.102±.055	0.385±.203
	CBO	0.424±.246	0.172±.124	0.175±.088	0.155±.072	0.367±.211
	COBOL	0.447±.270	0.171±.098	0.126±.072	0.033±.016	0.000±.000
	cCBO	0.235±.135	0.184±.119	0.309±.293	0.024±.017	0.004±.002
	DCBO	0.465±.288	0.036±.020	0.000±.000	0.163±.106	0.140±.091
	MCBO	0.071±.045	0.162±.103	0.100±.093	0.000±.000	0.100±.082
	HCBO	0.285±.162	0.107±.091	0.010±.005	0.146±.072	0.105±.072
	ECBO	**0.493±.281**	**0.299±.155**	**0.335±.147**	**0.176±.091**	**0.394±.203**
GAP	BO	0.346±.204	0.401±.172	0.310±.089	0.148±.082	0.817±.194
	CBO	0.810±.198	0.658±.104	0.577±.108	0.296±.109	0.692±.204
	COBOL	0.922±.125	0.593±.108	0.623±.031	0.431±.092	0.000±.000
	cCBO	0.730±.102	**0.691±.123**	**0.842±.132**	0.228±.076	0.501±.129
	DCBO	0.949±.213	0.482±.032	0.487±.004	0.448±.098	0.507±.066
	MCBO	0.547±.142	0.652±.132	0.510±.104	0.000±.000	0.527±.103
	HCBO	0.727±.136	0.559±.106	0.335±.032	**0.484±.132**	0.543±.032
	ECBO	**0.988±.124**	0.538±.151	0.684±.082	0.305±.192	**0.822±.195**

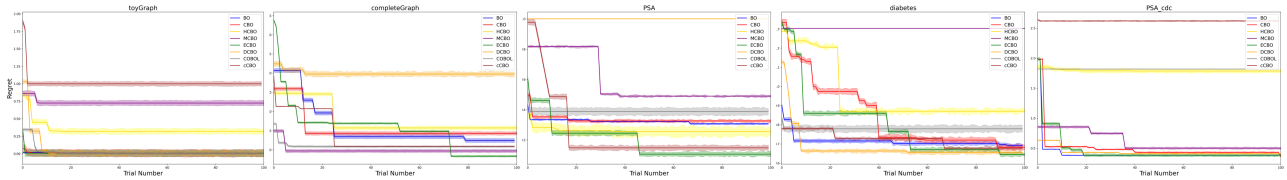

Figure 1: Best-so-far simple regret per trial visualization for eight models on five datasets. Solid lines show the mean; shaded regions show the standard deviation across twenty random seeds, including randomized initializations of the observation and intervention data.

observational data $\mathcal{D}_0^O$ and interventional data $\mathcal{D}_0^I$. We report the mean and standard deviation of GAP and PA-GAP in Table 1. We fix all algorithmic hyperparameters (see Appendix 1) to isolate the effect of fine-tuning. Each dataset uses 250 observational samples and 100 optimization trials. Budget-limited variants with 50 and 20 trials are reported in Appendix 9.

From Figure 1, we observe that ECBO continues to optimize throughout the entire process and is able to find the best optimal value across all datasets. Especially in CompleteGraph, ECBO starts with the worst initial point but is able to continue exploring and optimizing with the aid of the expert preference guidance, eventually smallest best-so-far simple regret by finding the closest point to the best theoretical value.

In Table 1, ECBO achieves the best PA-GAP scores, reflecting both the highest optimization efficiency and the best optimal values across synthetic and real-world scenarios. In particular, under realistic conditions, ECBO outperforms all baselines, including on the challenging Diabetes dataset, which has the largest number of nodes among all scenarios. DCBO, MCBO, and COBOL fail to make progress on PSA, Diabetes, and PSA_CDC, respectively, resulting in PA-GAP scores of 0, as they are unable to identify any solution better than the initial best observation. In terms of GAP, ECBO has no superior performance because the additive definition of GAP introduces a ranking bias, with which a trajectory that is overall superior may be judged as inferior. Detailed examples are in Appendix 8.2.

Table 2: Performance ECBO across different wrong SCMs. Values are PA-GAP (mean ± std. over 20 random initializations); Best values per dataset are in **bold**.

Variants	ToyGraph	PSA	Diabetes	PSA_CDC
Missing Edge	0.483±.182	0.305±.125	0.175±.073	0.392±.194
Missing node	0.456±.207	0.294±.142	0.172±.065	0.392±.182
Original	**0.493±.281**	**0.335±.147**	**0.176±.091**	**0.394±.203**

4.5 ECBO Robustness on Inaccurate SCMs

Even though we curated and validated the SCMs inferred by LiNGAM for Diabetes and PSA_CDC using domain knowledge and multiple model-fit tests [3], they may still not be perfectly accurate. To further evaluate robustness, we also tested model performance un-

[3]Model-fit tests details in the Appendix 7.3 and 7.4

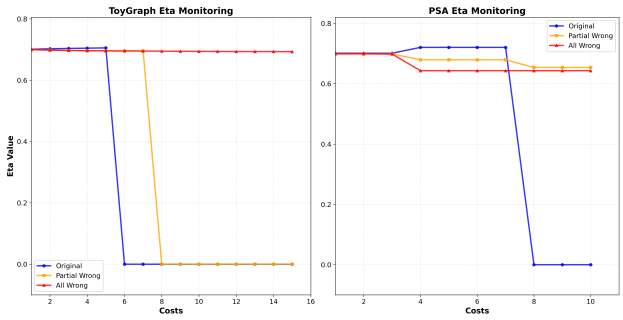

Figure 2: η monitoring on ToyGraph and PSA under three different preference settings.

Table 3: Ablation of fallback, handover, and η–update mechanisms. Values are PA-GAP (mean ± std. over 20 random intialization); Best values per dataset are in **bold**

Variant	ToyGraph	PSA	Diabetes
w/o fallback	0.492±.262	0.254±.033	0.146±.019
w/o handover	0.491±.228	0.258±.010	0.153±.028
w/o η–update	0.491±.283	0.276±.015	**0.178±.102**
Original	**0.493±.281**	**0.335±.147**	0.176±.091

der perturbed SCMs. In Table 2, the missing edge setting refers to the intentional removal of up to two edges from selected nodes, ensuring that the nodes remain connected to the target. The missing node setting refers to the removal of a node together with all of its incident edges. Visualizations of these perturbed SCMs are provided in Appendix 10.

We observed that model performance degraded more noticeably on ToyGraph and PSA, where the true SCMs are given. In contrast, for Diabetes and PSA_CDC, whose SCMs were inferred via LiNGAM, the effect was much smaller. This implies that performance degradation is more sensitive when the ground-truth SCM is available, but less impactful when the SCM is already approximate.

Ablation Study Table 3 reveals a clear pattern: on simpler tasks (like ToyGraph), removing fallback, handover, or trust updates barely moves the needle, but on more challenging benchmarks (PSA and Diabetes), each mechanism becomes essential. In particular, disabling the trust-update consistently worsens performance under noisy expert feedback, whereas handover offers additional robustness when uncertainty is high. Interestingly, in the Case of Diabetes, expert guidance is generally reliable, so freezing η accelerates convergence and slightly raises PA-GAP; however, disabling handover still hurts by cutting off useful cues too early.

4.6 η-Monitoring

We investigate how expert preference settings impact the trust coefficient η, which measures the degree of reliance on expert guidance. Because each intervention subset has its own expert surrogate, we track η over time for the subset tied to the final optimal intervention ($\mathbf{x}_t^*$): $\{Statin\}$ on PSA and (Z) on ToyGraph. We log η at each intervention on that subset; the x-axis in Fig. 2 is the cumulative intervention cost. We compare three configurations: (i) Original (ideal ECBO guidance), (ii) Partial-wrong (some preferences flipped), and (iii) All-wrong (all preferences contradict correct ECBO guidance). The expert references details for every dataset are included in Appendix 11.

On ToyGraph, correct guidance holds η at its initial 0.7 until cost 6, then triggers a clean handover ($\eta \to 0$), whereas Partial Wrong delays this drop until cost 8, reflecting occasional mislabels, and All Wrong never falls below its starting value, indicating no fallback under fully contradictory feedback. On PSA, the Original guidance again maintains η around 0.71 until cost 8 before handing over; in the wrong cases, η never rises and quickly settles at a low level, confirming that our safety rule prevents overtrust when expert guidance is misleading. These plots confirm that our revised trust-update and safety-fallback rules enable a timely handover under correct feedback, graceful trust withdrawal under partial misguidance, and a stable low-trust floor when expert advice is entirely erroneous.

5 Conclusion

In this work, we have introduced ECBO, a novel framework that integrates domain expert feedback into causal Bayesian optimization by augmenting the classic EI acquisition with an adaptive trust coefficient, gradient-aligned safety fallback, and an automated handover stopping criterion. Our method preserves the rigour of causal intervention through a context-aware Gaussian process posterior while enabling flexible, multi-categorical expert preferences and robustly managing adversarial or uncertain feedback. To further enhance scalability in high-uncertainty regimes, we propose an uncertainty handling rule that skips surrogate and trust-updates on purely uncertain labels, reducing query and computational overhead without degrading optimization quality. Extensive experiments on synthetic SCM benchmarks and real-world tasks demonstrate that ECBO outperforms vanilla EI in both solution quality and efficiency.

References

D. R. Jones, M. Schonlau, and W. J. Welch (1998). Efficient global optimization of expensive black-box functions. *Journal of Global Optimization* **13**(4):455–492. Springer.

V. Aglietti, X. Lu, A. Paleyes, and J. González (2020). Causal Bayesian optimization. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, 3155–3164. PMLR.

J. Pearl (2009). *Causality*. Cambridge: Cambridge University Press.

S. Sussex, A. Makarova, and A. Krause (2023). Model-based causal Bayesian optimization. In *Proceedings of the International Conference on Learning Representations (ICLR)*.

Y. Wu, W. Wang, Y. Zhang, M. Li, Y. Liu, H. Qian, and A. Zhou (2024). High-dimensional causal Bayesian optimization. In *Proceedings of the European Conference on Artificial Intelligence (ECAI)*.

W. Xu, M. Adachi, C. Jones, and M. A. Osborne (2024). Principled Bayesian optimization in collaboration with human experts. *Advances in Neural Information Processing Systems* **37**:104091–104137.

A. K. AV, S. Rana, A. Shilton, and S. Venkatesh (2022). Human-AI collaborative Bayesian optimisation. *Advances in Neural Information Processing Systems* **35**:16233–16245.

S. Gupta, A. Shilton, A. K. AV, S. Ryan, M. Abdolshah, H. Le, S. Rana, J. Berk, M. Rashid, and S. Venkatesh (2023). BO-Muse: A human expert and AI teaming framework for accelerated experimental design. *arXiv preprint arXiv:2303.01684*.

A. K. AV, A. Shilton, S. Gupta, S. Rana, S. Greenhill, and S. Venkatesh (2024). Enhanced Bayesian optimization via preferential modeling of abstract properties. *arXiv preprint arXiv:2402.17343*.

V. Aglietti, T. Damoulas, M. Álvarez, and J. González (2020). Multi-task causal learning with Gaussian processes. *Advances in Neural Information Processing Systems* **33**:6293–6304.

S. Sussex, P. G. Sessa, A. Makarova, and A. Krause (2024). Adversarial causal Bayesian optimization. In *Proceedings of the International Conference on Learning Representations (ICLR)*.

L. Gultchin, V. Aglietti, A. Bellot, and S. Chiappa (2023). Functional causal Bayesian optimization. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI)*, 756–765. PMLR.

V. Aglietti, N. Dhir, J. González, and T. Damoulas (2021). Dynamic causal Bayesian optimization. *Advances in Neural Information Processing Systems* **34**:10549–10560.

S. Shimizu (2014). LiNGAM: Non-Gaussian methods for estimating causal structures. *Behaviormetrika* **41**(1):65–98. Springer.

J. W. Smith, J. E. Everhart, W. C. Dickson, W. C. Knowler, and R. S. Johannes (1988). Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. In *Proceedings of the Annual Symposium on Computer Application in Medical Care*, 261.

Centers for Disease Control and Prevention (2011). National Health and Nutrition Examination Survey (NHANES) 2009–2010 data. `https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2009`. Accessed: 2025-07-27.

C. Thompson (2019). Causal graph analysis with the Causalgraph procedure. In *Proceedings of SAS Global Forum*.

S. Ren and X. Qian (2024). Causal Bayesian optimization via exogenous distribution learning. *arXiv preprint arXiv:2402.02277*.

V. Aglietti, A. Malek, I. Ktena, and S. Chiappa (2023). Constrained causal Bayesian optimization. In *Proceedings of the International Conference on Machine Learning (ICML)*, 304–321. PMLR.

J. Pearl *et al.* (2000). *Models, Reasoning and Inference*. Cambridge: Cambridge University Press.

M. A. Osborne, R. Garnett, and S. J. Roberts (2009). Gaussian processes for global optimization. In *Proceedings of the Third International Conference on Learning and Intelligent Optimization (LION3)*.

R. Garnett (2023). *Bayesian Optimization*. Cambridge: Cambridge University Press.

M. Adachi, B. Planden, D. A. Howey, M. A. Osborne, S. Orbell, N. Ares, K. Muandet, and S. L. Chau (2024). Looping in the human collaborative and explainable Bayesian optimization. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics (AISTATS)*. PMLR.

A. Cissé, X. Evangelopoulos, S. Carruthers, V. V. Gusev, and A. I. Cooper (2023). HypBO: Accelerating black-box scientific experiments using experts' hypotheses. *arXiv preprint arXiv:2308.11787*.

M. Adachi (2021). High-dimensional discrete Bayesian optimization with self-supervised representation learning for data-efficient materials exploration. In *NeurIPS 2021 Workshop on AI for Science*.

H. J. Kushner (1964). A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *Journal of Basic Engineering* **86**(1):97–106.

A. G. Zhilinskas (1975). Single-step Bayesian search method for an extremum of functions of a single variable. *Cybern Syst Anal* **11**:160–166.

J. Mockus (1974). On Bayesian methods for seeking the extremum. In *Optimization Techniques*.

Y. Wu, W. Wang, Y. Zhang, M. Li, Y. Liu, H. Qian, and A. Zhou (2024). High-dimensional causal Bayesian optimization. In *ECAI*, 2990–2997.

N. Branchini, V. Aglietti, N. Dhir, and T. Damoulas (2023). Causal entropy optimization. In F. Ruiz, J. Dy, and J.-W. van de Meent (eds.), *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, 8586–8605. PMLR.

S. Mukherjee, M. Zhang, S. Flaxman, and J. S. Vollmer (2024). Graph agnostic causal Bayesian optimisation. NeurIPS BDU Workshop 2024 Poster.

Z. Li, Y. Li, Z. Lian, and R. Zheng (2023). Improved causal Bayesian optimization algorithm with counter-noise acquisition function and supervised prior estimation. *Journal of Physics: Conference Series* **2547**(1):012017. IOP Publishing.

S. Sussex, P. G. Sessa, A. Makarova, and A. Krause (2024). Adversarial causal Bayesian optimization. In *Proceedings of the International Conference on Learning Representations (ICLR)*.

S. Bhatija, P.-D. Zuercher, J. Thumm, and T. Bohné (2025). Multi-Objective Causal Bayesian Optimization. In *Proceedings of the International Conference on Machine Learning (ICML)*, to appear. PMLR.

A. D. Bull (2011). Convergence rates of efficient global optimization algorithms. *Journal of Machine Learning Research*, 12(10).

Z. Li, Y. Li, Z. Lian, and R. Zheng (2023). Improved Causal Bayesian Optimization Algorithm with Counter-Noise Acquisition Function and Supervised Prior Estimation. In *Journal of Physics: Conference Series*, vol. 2547, no. 1, pp. 012017. IOP Publishing.

V. Arsenyan, A. Grosnit, and H. Bou-Ammar (2023). Contextual causal Bayesian optimisation. *arXiv preprint arXiv:2301.12412*.

Confidence-Based Handover for Causal Bayesian Optimization with Adaptive Expert Trust Supplementary Materials

A Training Settings

All experiments were carried out on a single desktop workstation with the following hardware and software configuration:

- **Operating System**: MacOS 15.5
- **CPU**: Apple M3 Max
- **System Memory**: 36GB
- **Libaries**: Python 3.8

The Table 4 includes hyperparameter values of ECBO used in all datasets.

Table 4: Hyperparameter values used in the experiment

Hyperparameter	Value
γ	0.05
η_0	0.7
α	0.7
β	0
g_{thr}	0.4

B Related Works

From BO to Causal BO. Classical Bayesian Optimization (BO) [1] constructs a probabilistic surrogate model, commonly a Gaussian Process (GP), along with an acquisition function such as Expected Improvement (EI) or Upper Confidence Bound (UCB) to efficiently search for the optimum of an expensive black-box objective. Causal Bayesian Optimization (CBO) [2] extends this framework by integrating Structural Causal Models (SCMs) and do-calculus into the BO paradigm. This elevates the notion of an "evaluation point" from an input (x) to an intervention ($\text{do}(X = x)$), enabling the search space to be pruned on the causal graph via concepts such as MIS and POMIS. Moreover, CBO leverages causal priors as GP means or structured kernels and employs *causalEI* with cost- or observation-intervention trade-offs to guide sampling. Together, these ideas establish a general framework for sample-efficient exploration and exploitation under causal graph constraints.

CBO with a known graph. The core CBO framework restricts actions to minimal or possibly-optimal intervention sets (MIS/POMIS), initializes causal GP surrogates with do-calculus priors, and uses causal expected improvement with cost and an observation–intervention trade-off. Constrained CBO (cCBO) [19] introduces cMIS pruning and feasibility-weighted cEI to satisfy safety constraints. High-dimensional CBO (HCBO) [29] replaces exhaustive MIS search with a submodular coverage heuristic that builds an efficient causal-coverage set (ECCIS) and compares subsets via a normalized UCB score. Multi-objective CBO (MO-CBO) [34] decomposes the search into local problems over possibly Pareto-optimal MIS and drives selection by relative hypervolume improvement. Dynamic CBO (DCBO) [13] transfers interventional information over time through a recursion on the graph and applies time-aware CEI.

Intervention-space and policy extensions. Model-based CBO (MCBO) [4] learns the full SCM by fitting a GP to each mechanism, propagates uncertainty through the graph, and selects actions via an optimism (UCB) rule with non-asymptotic regret bounds. Functional CBO (fCBO) [12] searches over policies $\pi : C \to X$ in an RKHS and uses functional EI. Contextual CBO (CoCaBO) [37] separates scope selection and within-scope optimization using a bandit layer (MAB-UCB) over mixed policy scopes and a BO inner loop. Adversarial CBO (ACBO) [33] casts CBO as a game and combines a multiplicative-weights update with a causal UCB oracle for no-regret learning under strategic or non-stationary interference.

Unknown causal structure. Causal Entropy Optimization (CEO) [30] maintains a posterior over graphs and chooses interventions by expected information gain on the optimizer, effectively performing Bayesian model averaging over SCMs. Graph-Agnostic CBO (GACBO) [31] maintains a confidence set of plausible graphs and mechanisms and applies an optimistic UCB over this set, providing a scalable alternative to entropy-based selection.

Robustness and priors. Noise-robust CBO [36] adjusts EI to counter measurement noise and initializes GP means with supervised priors learned from observational data, improving early-stage stability.

Positioning of ECBO. Complementary to the directions above, ECBO centers on expert–algorithm collaboration and safe handover.
(1) Unlike cCBO [19], which emphasizes hard-constraint feasibility, ECBO focuses on a risk-aware transition strategy that dynamically balances trust between expert advice and data-driven evidence.
(2) In contrast to fCBO [12] or CoCaBO [37], which pursue conditional or policy optimality, ECBO explicitly models when to follow expert guidance and when to switch to autonomous decision-making as part of the acquisition process, adapting over time with changing uncertainty and cost.
(3) Compared with MCBO [4] and ACBO [33], which stress optimistic uncertainty propagation or adversarial robustness, ECBO highlights a closed-loop update between expert priors and intervention feedback, producing interpretable trust curves and ensuring safe transitions.

Overall, ECBO unifies causal priors, safety and cost considerations, and dynamic expert trust within the Bayesian optimization loop, filling a crucial gap in the CBO landscape for human–machine collaborative decision-making.

C Preference Modeling Weight Calculation

According to main text Section 3.1, we have four types of preferences: **Exclude**, **Promote**, **Suppress** and **Uncertain**. Let us assume that their weights $\omega_k\left(\tau_k(x_j)\right)$ are:

- **Exclude**: 0.1
- **Promote**: 1.2
- **Suppress**: 0.8
- **Uncertain**: 1

In example of Diabetes, we have two interventional variables: Insulin and BloodPressue, which represents the amount of insulin the patient takes and the blood pressure of the patient. In order to suppress the target outcome, the possibility that the patient has diabetes, expert preference suggests to **Promote** Insulin and **Suppress** the BloodPressure. Therefore, in case of different intervention sets, the $w(\mathbf{x})$ will be:

- **Insulin**: 1.2
- **BloodPressure**: 0.8
- **Insulin, BloodPressure**: 1.2 * 0.8 = 0.96

D Proofs for §3.5

Proof of Lemma 1. Since co and w are bounded and $\eta_t \in [0,1]$, Eq. (14) implies

$$\frac{1+\eta_t(w_{\min}-1)}{c_{\max}}\,\mathrm{EI} \;\le\; \mathrm{ECEI} \;\le\; \frac{1+\eta_t(w_{\max}-1)}{c_{\min}}\,\mathrm{EI}. \qquad \square$$

Proof of Theorem 1. By the ρ_t-approximation,

$$\mathrm{ECEI}(x_t^{\mathrm{ec}}) \;\ge\; \rho_t \max_x \mathrm{ECEI}(x) \;\ge\; \rho_t\,\alpha_{\min,t} \max_x \mathrm{EI}(x).$$

Also $\mathrm{ECEI}(x_t^{\mathrm{ec}}) \le \alpha_{\max,t}\,\mathrm{EI}(x_t^{\mathrm{ec}})$. Thus

$$\mathrm{EI}(x_t^{\mathrm{ec}}) \;\ge\; \Big(\rho_t\,\alpha_{\min,t}/\alpha_{\max,t}\Big) \max_x \mathrm{EI}(x),$$

and replacing the maximum by x_t^{base} gives the claim. $\square$

Proof of Theorem 2. When $\eta_\tau = 0$, ECEI $\equiv$ EI/co. With a fixed candidate pool and fixed positive co, the maximizers of ECEI and EI coincide; with identical histories, the two policies agree for all $t \ge \tau$. $\square$

Proof of Theorem 3. Let $I_t^{\mathrm{base}} = \max_{x\in\mathcal{S}_t^{\mathrm{cand}}} \mathrm{EI}_t(x)$ and $I_t^{\mathrm{ec}} = \mathrm{EI}_t(x_t^{\mathrm{ec}})$. By Theorem 1, $I_t^{\mathrm{ec}} \ge \kappa_t I_t^{\mathrm{base}}$. If $\inf_t \kappa_t > 0$, or once the coupling of Theorem 2 applies, ECBO–EI's per-step improvement is of the same order as the baseline's, hence it inherits the best-so-far and cumulative regret rates. $\square$

E Complexity Analysis

Computational complexity. Let $M = |\mathcal{S}_{\mathrm{exp}}|$ and $n_S(t)$ be the number of interventional data used by GP $\mathcal{G}_S$ at round t. Each round optimizes ECEI^S for all S with a small inner budget $K_S(t)$ of acquisition evaluations. With a cached Cholesky, one GP posterior (and gradient) call is $O(n_S(t)^2)$. Acquisition over all subsets therefore costs $O\Big(\sum_{S\in\mathcal{S}_{\mathrm{exp}}} K_S(t)\,n_S(t)^2\Big)$ per round. We then refit only $\mathcal{G}_{S_t}$; exact GP training is $O(n_{S_t}(t{+}1)^3)$. Selecting the best tuple in $\mathcal{C}_t$ is $O(M)$. Hence the per-round time is $O\Big(\sum_S K_S(t)\,n_S(t)^2\Big) \;+\; O\big(n_{S_t}(t{+}1)^3\big)$, typically dominated by the cubic update when data accumulate. The expert weighting and safety checks add only lower-order overhead and do not change these rates.

Space complexity. We keep one GP per subset. Storing the Cholesky factor for $\mathcal{G}_S$ uses $O(n_S(t)^2)$ memory (inputs add $O(n_S(t))$ to constants). Total memory at round t is $O\Big(\sum_{S\in\mathcal{S}_{\mathrm{exp}}} n_S(t)^2\Big)$, dominated by the quadratic GP terms. If data are spread roughly evenly across M subsets, $n_S(t)\approx t/M$ and memory scales as $O(t^2/M)$.

F Baseline Details

- **BO** [2]: Constructs a probabilistic surrogate model, commonly a Gaussian Process (GP), along with an acquisition function
- **CBO** [2]: Extend the BO framework by integrating Structural Causal Models (SCMs) and the do-calculus.
- **MCBO** [4]: Learn the full SCM by fitting a GP to each mechanism, propagates uncertainty through the graph. MCBO supports both soft and hard interventions; we only consider its performance under the hard intervention setting to align with our experimental setup.
- **HCBO** [29]: Build an efficient causal-coverage set to support high dimension SCM optimization by exhaustive MIS search with a submodular coverage heuristic.
- **cCBO** [19]: Introduce cMIS pruning and feasibility-weighted cEI to satisfy safety constraints.
- **DCBO** [13]: Transfers interventional information over time through a recursion on the graph and applies CEI. Although it mainly addresses a dynamic SCM that changes over time, it also supports the fixed SCM optimization task.

Table 5: Summary of existing CBO works

Model	SCM	Intervention Type	Acquisition
CBO	fixed	hard	causal EI
MCBO	fixed	hard/soft	UCB
HCBO	fixed	hard	ISSF & UCB
fCBO	fixed	hard/soft	causal EI
cCBO	fixed	hard	constrained EI
ACBO	fixed	soft	UCB
DCBO	dynamic	hard	causal EI
ECBO	fixed	hard	expert-EI

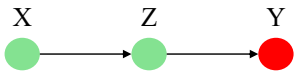

Figure 3: ToyGraph SCM. Green nodes presents interventional variables; red node is the target variable.

G Dataset

G.1 ToyGraph

A simple synthetic SCM [2] consisting of three chain-structured nodes: X, Z, and Y, where X and Z are designated as interventional variables, and Y is the target variable. Figure 3 depicts the generated SCM and detailed SCM illustrated below:

Exogenous Variable

$$X \text{ (input variable, exogenous)}$$

G.1.1 Endogenous Variables

$$Z = e^{-X} + \varepsilon_Z$$
$$Y = \cos Z \ - \ e^{-Z/20} + \varepsilon_Y$$

Each $\varepsilon_\bullet$ represents an independent noise term.

G.1.2 CompleteGraph

A more complex synthetic SCM [2] comprising seven observed nodes A, B, C, D, E, F, Y and two latent variables U_1 and U_2. Among the observed nodes, B, D, E are interventional variables, A, C, F are non-manipulative variables, and Y is the target one.

Figure 4 depicts the generated SCM and detailed SCM illustrated below:

Latent Variables

$$U_1 = \varepsilon_{YA} \sim \mathcal{N}(0, 1)$$
$$U_2 = \varepsilon_{YB} \sim \mathcal{N}(0, 1)$$

Exogenous Variable

$$F = \varepsilon_F \sim \mathcal{N}(0, 1)$$

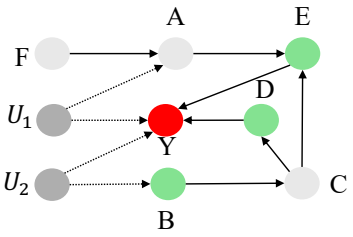

Figure 4: CompleteGraph SCM. Light-Grey nodes are nonintervention variables; dark-grey nodes are latent variables; green nodes presents interventional variables; red node is the target variable.

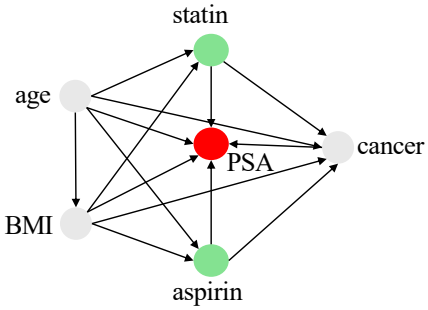

Figure 5: PSA SCM. Light-Grey nodes are nonintervention variables; green nodes presents interventional variables; red node is the target variable.

Endogenous Variables

$$
\begin{aligned}
A &= F^2 + U_1 + \varepsilon_A \\
B &= U_2 + \varepsilon_B \\
C &= e^{-B} + \varepsilon_C \\
D &= \frac{e^{-C}}{10} + \varepsilon_D \\
E &= \cos A + \frac{C}{10} + \varepsilon_E \\
Y &= \cos D + \sin E + U_1 + U_2\,\varepsilon_Y
\end{aligned}
$$

Each $\varepsilon_\bullet$ represents an independent noise term.

G.2 PSA

An SCM derived from a real-world healthcare setting [17], involving six observed variables: *Age*, *BMI*, *Aspirin*, *Statin*, *Cancer*, and *PSA*. The interventional variables in this setting are *Aspirin* and *Statin*, and the target variable is *PSA*.

Figure 5 depicts the generated SCM and detailed SCM illustrated below:

Exogenous Variable

$$\text{Age} \sim \text{Uniform}(55,\ 75)$$

Endogenous Variables

$$
\begin{aligned}
\mathrm{BMI} &= 27.0 \;-\; 0.01\,\mathrm{Age} \;+\; \varepsilon_{\mathrm{BMI}},\\
&\qquad \varepsilon_{\mathrm{BMI}} \sim \mathcal{N}(0,\, 0.7^2)\\
\mathrm{Aspirin} &= \sigma\big(-8.0 \;+\; 0.10\,\mathrm{Age} \;+\; 0.03\,\mathrm{BMI}\big) \;+\; \varepsilon_{\mathrm{Aspirin}}\\
\mathrm{Statin} &= \sigma\big(-13.0 \;+\; 0.10\,\mathrm{Age} \;+\; 0.20\,\mathrm{BMI}\big) \;+\; \varepsilon_{\mathrm{Statin}}\\
\mathrm{Cancer} &= \sigma\big(2.2 \;-\; 0.05\,\mathrm{Age} \;+\; 0.01\,\mathrm{BMI} \;-\; 0.04\,\mathrm{Statin}\\
&\qquad +\; 0.02\,\mathrm{Aspirin}\big) \;+\; \varepsilon_{\mathrm{Cancer}}\\
\mathrm{PSA} &= 6.8 \;+\; 0.04\,\mathrm{Age} \;-\; 0.15\,\mathrm{BMI}\\
&\qquad -\; 0.60\,\mathrm{Statin} \;+\; 0.55\,\mathrm{Aspirin} \;+\; 1.00\,\mathrm{Cancer}\\
&\qquad +\; \varepsilon_{\mathrm{PSA}},\\
&\qquad \varepsilon_{\mathrm{PSA}} \sim \mathcal{N}(0,\, 0.4^2)
\end{aligned}
$$

$\sigma(x) = \frac{1}{1+e^{-x}}$ denotes the sigmoid function, and each $\varepsilon_\bullet$ represents an independent noise term.

G.3 Diabetes

Records from a survey examination of Indian Health Service Hospital [15], including features of number of times pregancy (*Pregnancies*), plasma glucose concentration level (*Glucose*), diastolic blood pressure (*BloodPressure*), triceps skin fold thickness (*SkinThickness*), 2-hour serum insulin level (*Insulin*), body mass index (*BMI*), diabetes pedigree function (*DiabetesPedigreeFunction*), *Age* and *Outcome* indicating whether being diagnoised as diabetes. Finally, we determine *Insulin* and *BloodPressure* as interventional varibles since they both can be modulated by drugs.

In LiNGAM SCM generation [14], we define *Age* as exogenous node since it should be caused by other variables by forcing it has no incoming edges, and we determine *Outcome* as target variable by forcing it has no outgoing edges. The generated SCM passed Root Mean Square Error of Approximation (RMSEA) (< 0.05), Comparative Fit Index (CFI) (> 0.95) and Tucker-Lewis Index (TLI) (> 0.95) fitting tests:

- RMSEA: 030299558883832317
- CFI: 0.9893820750484923
- TLI: 0.9787641500969847

Figure 6 depicts the generated SCM and detailed SCM illustrated below:

G.3.1 Exogenous Variables

$$
\begin{aligned}
\mathrm{DiabetesPedigreeFunction} &\sim \mathrm{Uniform}(0.078,\; 2.42)\\
\mathrm{Age} &\sim \mathrm{Uniform}(21.0,\; 81.0)
\end{aligned}
$$

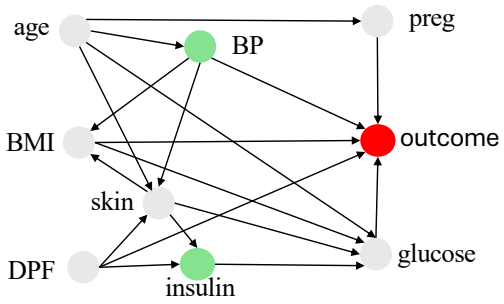

Figure 6: Diabetes SCM. Light-Grey nodes are nonintervention variables; green nodes presents interventional variables; red node is the target variable. *Skin* denotes *SkinThickness*, *BP* refers to *BloodPressure*, *Preg* stands for *Pregnancies*, and *DPF* is the abbreviation for *DiabetesPedigreeFunction*.

G.3.2 Endogenous Variables

$$
\begin{aligned}
\text{Pregnancies} &= -1.3394 + 0.1560 \cdot \text{Age} \\
\text{Glucose} &= 74.4170 - 0.2108 \cdot \text{SkinThickness} \\
&\quad + 0.0954 \cdot \text{Insulin} + 0.6635 \cdot \text{BMI} \\
&\quad + 0.6608 \cdot \text{Age} \\
\text{BloodPressure} &= 56.2742 + 0.3860 \cdot \text{Age} \\
\text{SkinThickness} &= 10.6921 + 0.1988 \cdot \text{BloodPressure} \\
&\quad + 8.5921 \cdot \text{DiabetesPedigreeFunction} \\
&\quad - 0.2390 \cdot \text{Age} \\
\text{Insulin} &= 3.2963 + 3.0020 \cdot \text{SkinThickness} \\
&\quad + 31.4762 \cdot \text{DiabetesPedigreeFunction} \\
\text{BMI} &= 23.2497 + 0.0767 \cdot \text{BloodPressure} \\
&\quad + 0.1677 \cdot \text{SkinThickness} \\
\text{Outcome} &= 1.7908 - 0.0215 \cdot \text{Pregnancies} \\
&\quad - 0.00586 \cdot \text{Glucose} \\
&\quad + 0.00091 \cdot \text{BloodPressure} \\
&\quad - 0.0114 \cdot \text{BMI} \\
&\quad - 0.0968 \cdot \text{DiabetesPedigreeFunction}
\end{aligned}
$$

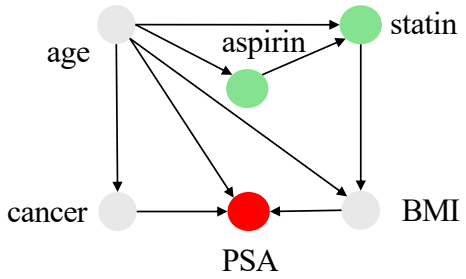

Figure 7: PSA-CDC SCM. Light-Grey nodes are nonintervention variables; green nodes presents interventional variables; red node is the target variable.

G.4 PSA-CDC

We collect and process the 2009–2010 Medical Conditions Questionnaire from the National Center for Health Statistics [16]. A total of 1,746 anonymized individual records were summarized and filtered, each originally has 8 features: *Sex*, *Race*, *PSA*, *Age*, *Aspirin*, *Statin*, *Cancer* and *BMI*. We drop the *Sex* and *Race* features when generating the SCM, as they negatively affect LiNGAM fitting. Following the approach used for the *Diabetes* dataset, we designate *Age* as the exogenous node and *PSA* as the target variable during SCM synthesis. The resulting SCM passes standard goodness-of-fit tests, including RMSEA (< 0.05), CFI (> 0.95), and TLI (> 0.95):

- RMSEA: 0.02918509041898147
- CFI: 0.9834272963873766
- TLI: 0.9585682409684413

Finally, we determine *Aspirin* and *Statin* as the interventional variables. Figure 7 depicts the generated SCM and detailed SCM illustrated below:

G.4.1 Exogenous Variables

$$\text{DiabetesPedigreeFunction} \sim \text{Uniform}(0.078,\ 2.42)$$
$$\text{Age} \sim \text{Uniform}(21.0,\ 81.0)$$

G.4.2 Endogenous Variables

$$
\begin{aligned}
\text{Pregnancies} &= -1.3394 + 0.156 \cdot \text{Age} \\
\text{Glucose} &= 74.4170 - 0.2108 \cdot \text{SkinThickness} \\
&\quad + 0.0954 \cdot \text{Insulin} + 0.6635 \cdot \text{BMI} \\
&\quad + 0.6608 \cdot \text{Age} \\
\text{BloodPressure} &= 56.2742 + 0.3860 \cdot \text{Age} \\
\text{SkinThickness} &= 10.6921 + 0.1988 \cdot \text{BloodPressure} \\
&\quad + 8.5921 \cdot \text{DiabetesPedigreeFunction} \\
&\quad - 0.2390 \cdot \text{Age} \\
\text{Insulin} &= 3.2963 + 3.0020 \cdot \text{SkinThickness} \\
&\quad + 31.4762 \cdot \text{DiabetesPedigreeFunction} \\
\text{BMI} &= 23.2497 + 0.0767 \cdot \text{BloodPressure} \\
&\quad + 0.1677 \cdot \text{SkinThickness} \\
\text{Outcome} &= 1.7908 - 0.0215 \cdot \text{Pregnancies} \\
&\quad - 0.00586 \cdot \text{Glucose} \\
&\quad + 0.00091 \cdot \text{BloodPressure} \\
&\quad - 0.0114 \cdot \text{BMI} \\
&\quad - 0.0968 \cdot \text{DiabetesPedigreeFunction}
\end{aligned}
$$

H Metrics Analysis

The GAP metric was introduced in DCBO [13] to quantify optimization efficiency. The GAP metric was defined as:

$$
GAP = \left[\underbrace{\frac{y(\mathbf{x}_t^*) - y(\mathbf{x}_{\text{init}})}{y^* - y(\mathbf{x}_{\text{init}})}}_{\text{Improvement Term}} + \underbrace{\frac{T - t^*}{T}}_{\text{Efficiency Term}} \right] \Bigg/ \underbrace{\left(1 + \frac{T-1}{T}\right)}_{\text{Normalization}}, \tag{1}
$$

where $y(\mathbf{x}) = \mathbb{E}[Y \mid do(\mathbf{X} = \mathbf{x})]$ denotes the objective function, y^* is the global optimum, and $\mathbf{x}_{\text{init}}$ and $\mathbf{x}_t^*$ are the initial and best-evaluated points at trial t, respectively. T is the total number of trials, and t^* denotes the trial index at which $y(\mathbf{x}_t^*)$ is discovered.

H.1 GAP Value Range

In DCBO, it defines that $t^* = 1$, so efficiency term equals $\frac{T-1}{T}$ if the algorithm converges at the first trial, which bound the GAP's upper limit as 1. However, it also defines the $t^* = 0$ when the algorithm is not converged. This actually makes the efficiency term equal to 1 with improvement term equal to 0, making the GAP lower bound become $1 \Big/ \left(1 + \frac{T-1}{T}\right) = \frac{T}{2T-1}$. Instead, if $t^* = T$ when the algorithm does not converge, both efficiency term and improvement term will equal to 0, which correctly establish the GAP value range $[0, 1]$

H.2 Comparison between GAP and PA-GAP

Although the GAP formulation appears reasonable, its limitation becomes clear through simple examples.

In Figure 8, both models start from the same initial optimum of 10 and discover an improved value of 5 at trial 10. Model 1 makes no further progress after trial 10. Model 2 finds the global optimum of 1 at trial 30. With $y^* = 1$ and $T = 40$, the original GAP metric gives Model 1 a score of 0.661 for $y(\mathbf{x}_t^*) = 5$ at $t^* = 10$, while Model 2 receives a lower score of 0.632 for $y(\mathbf{x}_t^*) = 1$ at $t^* = 30$. This outcome is due to ranking bias: both achieve the same intermediate optimum of 5 at the same trial with the same evaluation budget, yet Model 2, despite reaching a better final solution, is penalized.

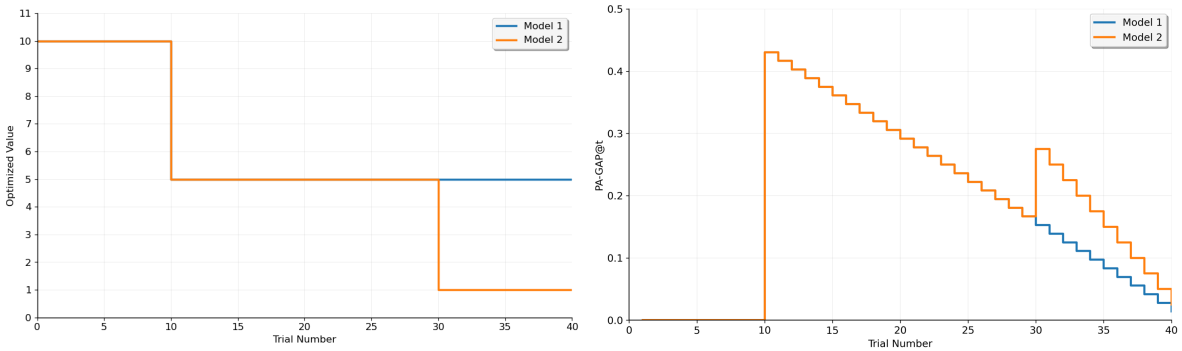

Figure 8: Demo experiment result comparison of optimization progress between two models

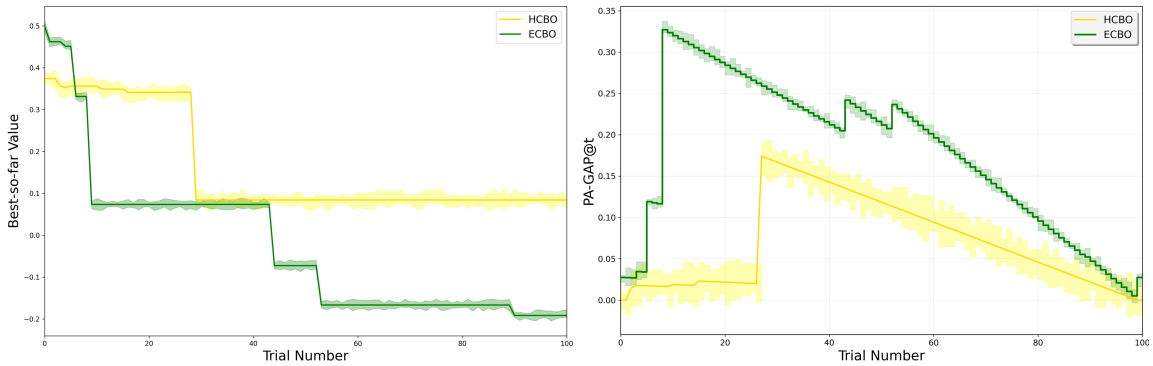

Figure 9: Diabetes experiment result comparison of optimization progress between two models

The ranking bias arises from two sources: the weak coupling between the improvement term and the efficiency term, and the dependence on only the final result instead of the entire optimization process.

To address this, we replace addition with multiplication to strengthen the link between improvement and efficiency, making it express in a way of using efficiency term to regularize the improvement term. It also allows us to removes the need for normalization term and simplifies the formulation. Furthermore, we incorporate averaging across all rounds to capture the full optimization process instead of evaluating solely on the final outcome. Finally, we introduce our Path-Aware GAP (PA-GAP):

$$\text{PA-GAP} = \frac{1}{T}\sum_{t=1}^{T}\left(\underbrace{\frac{y(\mathbf{x}_t^*) - y(\mathbf{x}_{\text{init}})}{y^* - y(\mathbf{x}_{\text{init}})}}_{\text{Improvement Term}} * \underbrace{\frac{T-(t-1)}{T}}_{\text{Efficiency Term}}\right). \tag{2}$$

The new metric constrains the improvement term to $[0, 1]$ and the efficiency term to $\left[\frac{1}{T}, 1\right]$. In the ideal scenario where the global optimum is reached at trial $t = 1$ with $y(\mathbf{x}_1^*) = y^*$, the score attains its maximum of 1. Conversely, if no improvement occurs throughout the entire optimization, the score will be 0.

Returning to the example in of Figure 8, the per-trial scores PA-GAP@t for both models are shown in right plot of Figure 8. PA-GAP@t captures the late improvement of Model 2 at $t = 30$. Summing PA-GAP@t over all t and averaging yields $\text{PA-GAP}_{\text{Model 1}} = 0.172$ and $\text{PA-GAP}_{\text{Model 2}} = 0.191$. Thus, PA-GAP mitigates the original GAP metric's bias against late-stage improvements by taking in account of the entire optimization trajectory.

In the Diabetes optimization experiment, ECBO improved from the initial value of 0.498 to -0.191 around trial 90, while HCBO converged from 0.374 to 0.084 around trial 28. With the best theoretical value at -0.836, the GAP score of HCBO is 0.484, compared to 0.305 for ECBO. Although ECBO shows a strong improvement score, its efficiency score is low. Notably, ECBO discovered a suboptimal value of 0.073 as early as trial 9. If

Table 6: Performance across different trial limits on different datasets. Each value is average ± std. across 20 random seeds and initialization of the observation and intervention data. Best values per dataset are in **bold**.

Trial Limit	Metrics	Model	ToyGraph	CompleteGraph	PSA	Diabetes	PSA_CDC
100	PA-GAP	BO	0.224±.139	0.198±.155	0.125±.064	0.102±.055	0.385±.203
		CBO	0.424±.246	0.172±.124	0.175±.088	0.155±.072	0.367±.211
		COBOL	0.447±.270	0.171±.098	0.126±.072	0.033±.016	0.000±.000
		cCBO	0.235±.135	0.184±.119	0.309±.293	0.024±.017	0.004±.002
		DCBO	0.465±.288	0.036±.020	0.000±.000	0.163±.106	0.140±.091
		MCBO	0.071±.045	0.162±.103	0.100±.093	0.000±.000	0.100±.082
		HCBO	0.285±.162	0.107±.091	0.010±.005	0.146±.072	0.105±.072
		ECBO	**0.493±.281**	**0.299±.155**	**0.335±.147**	**0.176±.091**	**0.394±.203**
	GAP	BO	0.346±.204	0.401±.172	0.310±.089	0.148±.082	0.817±.194
		CBO	0.810±.198	0.658±.104	0.577±.108	0.296±.109	0.692±.204
		COBOL	0.922±.125	0.593±.108	0.623±.031	0.431±.092	0.000±.000
		cCBO	0.730±.102	**0.691±.123**	**0.842±.132**	0.228±.076	0.501±.129
		DCBO	0.949±.213	0.482±.032	0.487±.004	0.448±.098	0.507±.066
		MCBO	0.547±.142	0.652±.132	0.510±.104	0.000±.000	0.527±.103
		HCBO	0.727±.136	0.559±.106	0.335±.032	**0.484±.132**	0.543±.032
		ECBO	**0.988±.124**	0.538±.151	0.684±.082	0.305±.192	**0.822±.195**
50	PA-GAP	BO	0.177±.129	0.130±.095	0.120±.064	0.096±.053	0.317±.179
		CBO	0.421±.238	0.131±.113	0.131±.082	0.109±.049	0.349±.209
		COBOL	0.407±.259	0.167±.094	0.127±.078	0.034±.106	0.000±.000
		cCBO	0.233±.134	0.191±.074	0.263±.170	0.011±.012	0.004±.002
		DCBO	0.410±.288	0.033±.018	0.000±.000	0.180±.101	0.110±.084
		MCBO	0.063±.044	0.150±.099	0.034±.049	0.000±.000	0.033±.034
		HCBO	0.262±.147	0.050±.060	0.009±.005	0.136±.046	0.063±.046
		ECBO	**0.498±.281**	**0.270±.145**	**0.266±.128**	**0.137±.086**	**0.368±.179**
	GAP	BO	0.296±.192	0.531±.082	0.366±.086	0.143±.032	**0.809±.146**
		CBO	0.676±.201	0.597±.108	0.598±.087	0.321±182	0.487±.172
		COBOL	0.853±.145	0.508±.119	0.621±.032	0.528±.089	0.000±.000
		cCBO	0.723±.098	0.565±.103	**0.758±.121**	0.178±.052	0.498±.031
		DCBO	0.896±.183	0.422±.012	0.474±.002	**0.630±.078**	0.332±.032
		MCBO	0.513±.106	**0.630±.099**	0.315±.087	0.000±.000	0.342±.098
		HCBO	0.643±.132	0.432±.087	0.155±.010	0.304±.102	0.421±.016
		ECBO	**0.988±.121**	0.578±.125	0.448±.112	0.267±.153	0.717±.103
20	PA-GAP	BO	0.088±.096	0.026±.035	0.121±.064	0.085±.042	**0.365±.148**
		CBO	0.392±.228	0.034±.053	0.034±.081	0.067±.033	0.311±.208
		COBOL	0.298±.227	0.155±.087	0.132±.072	0.035±.106	0.000±.000
		cCBO	0.227±.131	0.147±.077	0.143±.030	0.000±.000	0.004±.002
		DCBO	**0.526±.288**	0.024±.013	0.000±.000	**0.145±.077**	0.048±.057
		MCBO	0.040±.037	0.116±.092	0.000±.000	0.000±.000	0.000±.000
		HCBO	0.200±.177	0.000±.001	0.008±.004	0.121±.010	0.017±.010
		ECBO	0.513±.281	**0.223±.107**	**0.182±.069**	0.085±.057	0.290±142
	GAP	BO	0.383±.098	0.239±.062	0.604±.069	0.485±.042	0.657±.163
		CBO	0.510±198	0.392±.089	0.392±.092	0.246±.125	**0.809±.163**
		COBOL	0.636±.178	0.235±.072	**0.615±.025**	**0.521±.082**	0.000±.000
		cCBO	0.702±.103	0.367±.098	0.490±.083	0.000±.000	0.490±.025
		DCBO	0.729±.172	0.231±.011	0.432±.001	0.508±.067	0.406±.032
		MCBO	0.406±.089	0.558±.072	0.000±.000	0.000±.000	0.000±.000
		HCBO	0.372±.098	0.166±.010	0.172±.003	0.314±.052	0.113±.009
		ECBO	**0.987±.102**	**0.559±.107**	0.550±.108	0.453±.132	0.414±.104

optimization had stopped there, ECBO would achieve a higher GAP score of 0.617.

In contrast, PA-GAP computes the average PA-GAP@t across all trials, thereby reflecting the entire optimization trajectory. Under this metric, ECBO achieves a PA-GAP of 0.176, while HCBO scores 0.146.

I Extended Experiment

Table 6 extends the experiment to different trial limits to simulate performance under varying budget constraints. ECBO achieves the best PA-GAP performance with trial limits of 100 and 50. However, BO and DCBO outperform when the limit is 20, indicating faster convergence and an advantage under tight budgets. ECBO shows advantages in GAP score due to its continued optimization, which raises the improvement term but lowers efficiency, resulting in reduced overall GAP despite sustained progress.

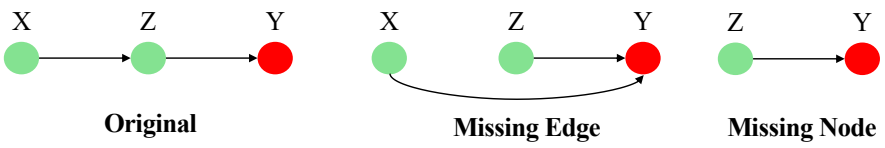

Figure 10: Wrong SCMs of ToyGraph we used

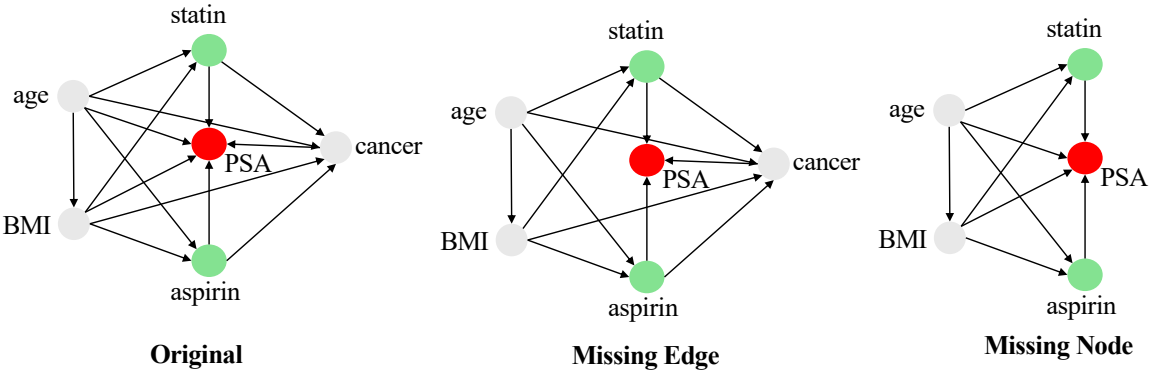

Figure 11: Wrong SCMs of PSA we used

J Wrong SCMs

In Section 4.5, we manually constructed incorrect SCMs for four datasets: ToyGraph, PSA, Diabetes and PSA_CDC. Each dataset includes two types of incorrect SCMs: **Missing Edge** and **Missing Node**.

Figure 10 illustrates the ToyGraph dataset, where the **Original** label denotes the correct SCM. In the **Missing Edge** version, the edge $X \to Z$ is removed and a compensatory edge $X \to Y$ is added to ensure node X remains connected. In the **Missing Node** version, node X and its outgoing edge $X \to Z$ are completely removed.

Figure 11 illustrates the PSA dataset, where the **Original** label denotes the correct SCM. In the **Missing Edge** version, the edge $age \to PSA$ and $BMI \to PSA$. In the **Missing Node** version, node *cancer* and its connected edges are completely removed.

Figure 12 illustrates the Diabetes dataset, where the **Original** label denotes the correct SCM. In the **Missing Edge** version, the edge $age \to glucose$ and $BMI \to glucose$. In the **Missing Node** version, node *preg* and its connected edges are completely removed.

Figure 13 illustrates the PSA_CDC dataset, where the **Original** label denotes the correct SCM. In the **Missing Edge** version, the edge $age \to BMI$ and $age \to statin$. In the **Missing Node** version, node *cancer* and its connected edges are completely removed.

K Expert Preference

Table 7 summarizes the expert preferences used for ECBO.

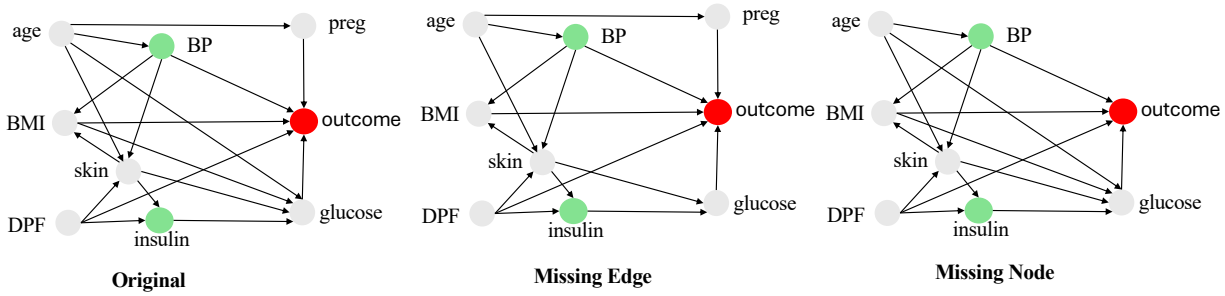

Figure 12: Wrong SCMs of Diabetes we used

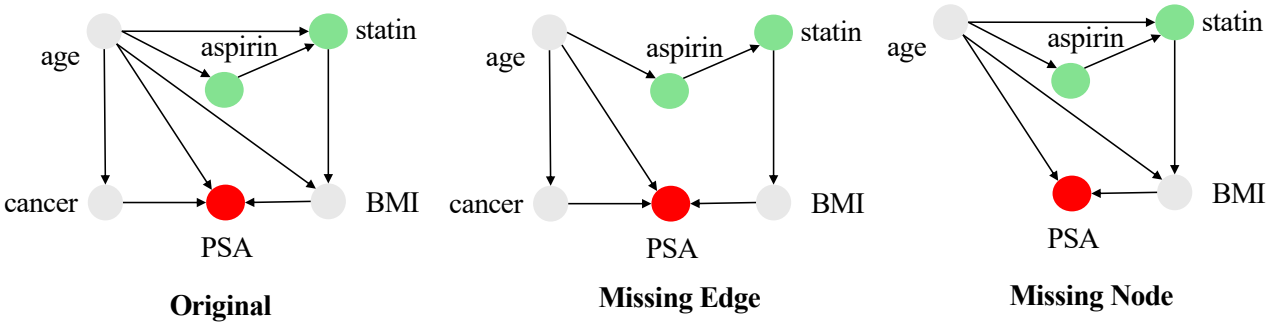

Figure 13: Wrong SCMs of PSA_CDC we used

Table 7: Expert-preference settings for each dataset.

Dataset	Variable	ECBO
ToyGraph	X	exclude
	Z	promote
CompleteGraph	B	exclude
	D	uncertain
	E	uncertain
PSA	Aspirin	promote
	Statin	uncertain
Diabetes	Insulin	promote
	BloodPressure	suppress
PSA–CDC	Aspirin	promote
	Statin	suppress