Learning the Best Under Constraints: A Duality-Based Framework

Mingjie Hu[1,2] Enlu Zhou[2] Jianqiang Hu[1]

[1]School of Management, Fudan University
[2]H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology

ICML 2026

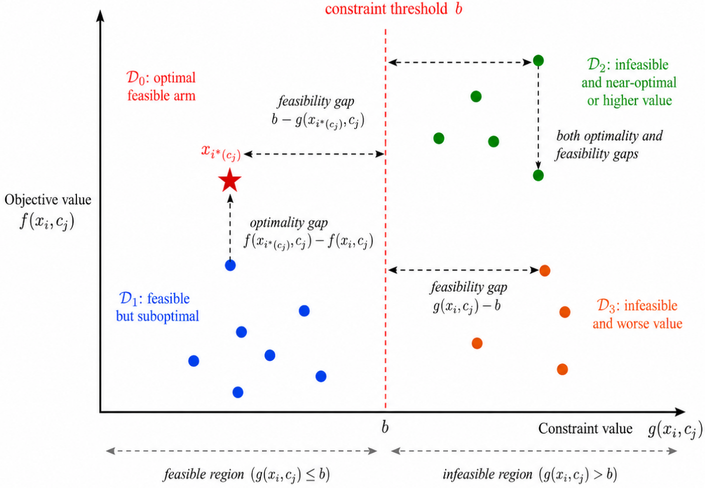
July 6, 2026

Motivation

- **Best Arm Identification (BAI):** identify the best arm with high confidence.
- **Our setting:** constrained linear BAI with covariate selection.
  - Multiple performance metrics.
  - Means depend on both arms and covariates.
  - Sampling is over arm-covariate pairs.
- **Applications:**
  - Personalized medicine: efficacy vs. side effects.
  - Inventory management: revenue vs. lead time and satisfaction.
- **Key challenge:** balance optimality and feasibility while actively selecting covariates.

How can we identify the best feasible arm for each covariate with as few samples as possible?

Constrained BAI with Covariate Selection



Main Result: Sample Complexity Lower Bound

Theorem: information-theoretic lower bound

Any δ-PAC algorithm must satisfy

$$\mathbb{E}[\tau] \geq \mathcal{H}^*(\mathcal{P})\, kl(\delta, 1-\delta), \qquad \liminf_{\delta \to 0} \frac{\mathbb{E}[\tau]}{\log(1/\delta)} \geq \mathcal{H}^*(\mathcal{P}).$$

$$\mathcal{H}^*(\mathcal{P})^{-1} = \max_{\omega \in \Omega} \min_{c_j \in \mathcal{C}} \Gamma(\omega, c_j, \mathcal{P}).$$

$$\Gamma(\omega, c_j, \mathcal{P}) = \frac{1}{2} \min \left\{ \begin{array}{l} \min_{x_i \neq x_{i^*(c_j)}} \left[\frac{\left((\phi(x_{i^*(c_j)}, c_j) - \phi(x_i, c_j))^\top \theta \right)^2}{\left\| \phi(x_{i^*(c_j)}, c_j) - \phi(x_i, c_j) \right\|^2_{\Lambda(\omega)^{-1}}} \mathbb{I}(x_i \in \mathcal{D}_1(c_j) \cup \mathcal{D}_3(c_j)) \right. \\ \qquad \left. + \frac{(b - \beta^\top \phi(x_i, c_j))^2}{\left\| \phi(x_i, c_j) \right\|^2_{\Lambda(\omega)^{-1}}} \mathbb{I}(x_i \in \mathcal{D}_2(c_j) \cup \mathcal{D}_3(c_j)) \right], \\ \frac{\left(b - \beta^\top \phi(x_{i^*(c_j)}, c_j) \right)^2}{\left\| \phi(x_{i^*(c_j)}, c_j) \right\|^2_{\Lambda(\omega)^{-1}}} \end{array} \right\}.$$

Optimality term Feasibility term

Duality Perspective

Key Message

The original lower-bound problem is hard; we solve a tractable dual of its surrogate problem.

- **Surrogate problem:**

$$\max_{\omega \in \Omega} \min_{c_j \in \mathcal{C}} \Gamma^s(\omega, c_j, \mathcal{P}).$$

- **Dual reformulation:**

$$\min_{\lambda} \quad \mathcal{Q}(\lambda, \mathcal{P}) = -\sum_{h \in [D]} \sqrt{\sum_{i \in [K], j \in [M]} \lambda_{ij} \chi_h(x_i, c_j)}$$

$$\text{s.t.} \quad \sum_{i \in [K], j \in [M]} \lambda_{ij} = 1, \quad \lambda_{ij} \geq 0.$$

The dual problem is simplex-constrained and easy to solve.

Duality-Based Decomposition Algorithm

Key Idea

Avoid solving the full surrogate problem at every iteration; update the dual variables by one cheap decomposition step.

- **Estimate the instance:** update $\hat{\mathcal{P}}(t)$, $\hat{\theta}(t)$, and $\hat{\beta}(t)$ from sampled data.
- **Dual update:** select one positive coordinate $(m(t), n(t))$ and search over reduced directions $\mathcal{D}^{m(t),n(t)}(\lambda(t-1))$.
- **One-step gradient descent:** update

$$\lambda(t) = \lambda(t-1) + s(t)d(t),$$

  only when sufficient descent is achieved.
- **Recover sampling ratio:** compute $\gamma(\hat{\mathcal{P}}(t))$ from the updated $\lambda(t)$.
- **Sampling rule:** sample the most under-sampled design point relative to $\gamma(\hat{\mathcal{P}}(t))$.

Algorithmic Guarantee

Why This Helps

Low per-iteration cost $+$ adaptive sampling $+$ relaxed optimality

- Avoids solving the full multi-level problem at every iteration.
- Uses one-step dual decomposition to update $\lambda(t)$ efficiently.
- Recovers the sampling ratio $\gamma(\hat{\mathcal{P}}(t))$ from the dual solution.

Theorem

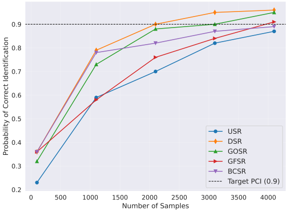
Under mild assumptions, the duality-based decomposition algorithm is δ-PAC and satisfies

$$\mathbb{P}\left(\limsup_{\delta\to 0}\frac{\tau}{\log(1/\delta)} \le \mathcal{U}^*(\mathcal{P})\right) = 1,$$

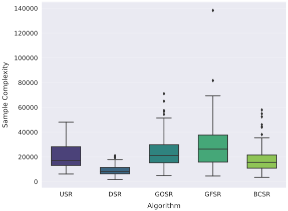
and

$$\limsup_{\delta\to 0}\frac{\mathbb{E}[\tau]}{\log(1/\delta)} \lesssim \mathcal{U}^*(\mathcal{P}).$$

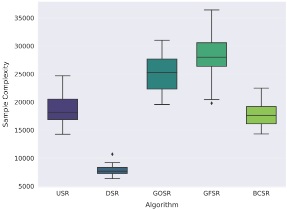
Numerical Results



PCI



Sample complexity



Robustness under t-noise

The proposed duality-based decomposition algorithm achieves better sampling efficiency while maintaining reliable identification performance.