MetaGameBO: Hierarchical Game-Theoretic Driven Robust Meta-Learning for Bayesian Optimization

Hui Li[1,2,3], **Huafeng Liu**[2,3], **Yiran Fu**[1,2,3], **Shuyang Lin**[1,2,3], **Baoxin Zhang**[1,2,3], **Deqiang Ouyang**[4,*], **Liping Jing**[1,2,3,*], **Jian Yu**[1,2,3]

[1]State Key Laboratory of Advanced Rail Autonomous Operation, Beijing, 100044, China
[2]School of Computer Science and Technology, Beijing Jiaotong University, Beijing, 100044, China
[3]Beijing Key Laboratory of Traffic Data Mining and Embodied Intelligence, Beijing, 100044, China
[4]Collage of Computer Science, Chongqing University, Chongqing, 400044, China
{huili97, hfliu1, yiranfu, sylin1, baoxin, lpjing, jianyu}@bjtu.edu.cn, deqiangouyang@cqu.edu.cn

Abstract

Meta-learning for Bayesian optimization accelerates optimization by leveraging knowledge from previous tasks, but existing methods optimize for average performance and fail on challenging outlier tasks critical in practice. These limitations become particularly severe when target tasks exhibit distribution shifts or when optimization budgets are limited in real-world applications. We introduce MetaGameBO, a hierarchical game-theoretic framework that formulates meta-learning as robust optimization through CVaR-based task selection and diversity-aware sample learning. Our approach incorporates uncertainty-aware adaptation via probabilistic embeddings and Thompson sampling for robust generalization to out-of-distribution targets. We establish theoretical guarantees including convergence to game-theoretic equilibria and improved sample complexity, and demonstrate substantial improvements with 95.7% reduction in average loss and 88.6% lower tail risk compared to state-of-the-art methods on challenging tasks and distribution shifts.

Introduction

Consider a critical scenario in automated machine learning: production systems need to optimize hyperparameters across diverse datasets, ranging from curated benchmarks to noisy real-world streams often containing extreme outliers. Traditional Bayesian optimization (BO) methods, while effective on single tasks, falter when requiring cross-task knowledge transfer or facing challenging scenarios that constitute the performance bottleneck. This fundamental limitation has driven the development of meta-learning approaches for BO. However, a critical limitation of current methods remains: their optimization focus on average-case performance often comes at the expense of robustness to worst-case scenarios.

Bayesian optimization is widely recognized as the gold standard for optimizing expensive black-box functions. Its applications span diverse domains, including hyperparameter tuning (Bergstra et al. 2011; Yu and Zhu 2020; Snoek, Larochelle, and Adams 2012), neural architecture search (Kandasamy et al. 2018; Jin, Song, and Hu 2019), and material design (Frazier and Wang 2016; Griffiths and Hernández-Lobato 2020). A core limitation of traditional BO approaches is their treatment of each optimization task in isolation. This neglects valuable knowledge from related tasks, resulting in inefficient exploration—especially detrimental under tight evaluation budgets (Feurer, Springenberg, and Hutter 2015).

Meta-learning for Bayesian optimization addresses this limitation by leveraging prior task experiences to accelerate optimization on new targets (Vanschoren 2018; Hospedales et al. 2021), with recent approaches demonstrating significant gains in sample efficiency and overall performance (Feurer, Letham, and Bakshy 2018; Wistuba, Schilling, and Schmidt-Thieme 2018; Volpp et al. 2018; Perrone et al. 2018). However, a fundamental challenge persists in practice: While existing methods perform well on easier tasks, they exhibit severe performance degradation in challenging scenarios, specifically those that often dictate practical system reliability.

This challenge stems from the inherent heterogeneity of real-world task distributions. As shown in Figure 1(b), optimization tasks exhibit varying risk profiles, with high-risk scenarios occurring less frequently but demanding careful handling to prevent catastrophic failures. Crucially, existing meta-learning BO methods suffer from three critical limitations: (1) Average-case bias: The dominant paradigm optimizes for average performance across the meta-training tasks. Consequently, these methods are vulnerable to failure on challenging outlier tasks, which can be mission-critical (Finn, Xu, and Levine 2018; Wang and Arora 2024). (2) Distribution shift vulnerability: When target tasks differ significantly from meta-training distributions, substantial performance degradation occurs (Finn et al. 2019; Jamal and Qi 2019).(3) Lack of robustness principles: Current approaches lack principled mechanisms to ensure worst-case performance while maintaining competitive average-case results.

Recent advances have begun addressing these challenges. For instance, MALIBO (Pan et al. 2024) employs a likelihood-free approach that enhances robustness to scale and noise variations. However, like most methods, it still targets average-case optimization and lacks explicit mechanisms for worst-case scenarios. On the theoretical front, game-theoretic frameworks for robust (minimax) meta-

*The corresponding authors

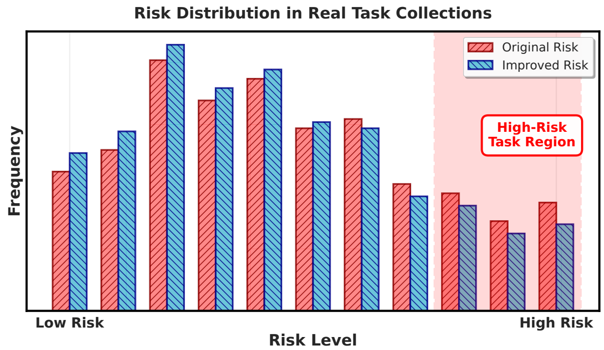

Figure 1: Motivation for MetaGameBO: Real-world task (Forrester function) collections exhibit heterogeneous risk distributions with critical high-risk tasks that occur less frequently but require careful handling.

learning (Yin et al. 2020; Lv et al. 2024) have been proposed. Yet, effectively adapting these frameworks to the sequential decision-making nature of Bayesian optimization, where sample efficiency constraints are critical, remains an open challenge.

To bridge this critical gap, we propose MetaGameBO, a hierarchical game-theoretic meta-learning framework designed to explicitly optimize robustness in challenging scenarios while maintaining high efficiency on typical tasks. Our key insight is to formulate meta-learning as a strategic interaction between two complementary mechanisms: task-level robustness through Conditional Value-at-Risk (CVaR) based selection in a Stackelberg game, and sample-level efficiency through diversity-aware weighted classification. This hierarchical design ensures robust performance across the heterogeneous risk spectrum shown in Figure 1, particularly excelling in the critical high-risk task region. Our main contributions include:

- **Hierarchical Game-Theoretic Framework**: A novel formulation of meta-learning BO as a hierarchical robust optimization problem, combining Stackelberg games for CVaR-based task selection with weighted classification for sample-level learning.
- **Uncertainty-Aware Adaptation**: Principled mechanisms for target task adaptation that leverage probabilistic task embeddings, uncertainty quantification (via Thompson sampling), and residual modeling (via gradient boosting) to maintain robustness under distribution shifts.
- **Theoretical and Empirical Validation**: Convergence guarantees, generalization bounds, and sample complexity results with logarithmic improvement over standard BO, validated through extensive experiments showing substantial performance gains on challenging tasks.

Related Work

This section reviews the key research areas that inform our approach: meta-learning foundations for Bayesian optimization, and robust optimization principles with game-theoretic frameworks for handling heterogeneous task distributions.

Meta-Learning for Bayesian Optimization

Bayesian optimization (BO) employs Gaussian Process surrogates with acquisition functions like Expected Improvement (Jones, Schonlau, and Welch 1998) and Upper Confidence Bound (Srinivas et al. 2010) to optimize expensive black-box functions (Shahriari et al. 2015; Brochu, Cora, and De Freitas 2010). Traditional BO methods treat each task in isolation (Snoek, Larochelle, and Adams 2012; Feurer, Springenberg, and Hutter 2015), leading to inefficient exploration when related tasks are available. Meta-learning approaches address this limitation by leveraging knowledge from previous tasks to accelerate optimization on new targets, ranging from warm-starting strategies (Feurer, Springenberg, and Hutter 2015; Wistuba, Schilling, and Schmidt-Thieme 2018) and transfer learning techniques (Perrone et al. 2018; Bardenet et al. 2013) to learned acquisition functions (Volpp et al. 2018; Zhou, Ma, and Blaschko 2024) and neural surrogates (Müller et al. 2023; Rothfuss et al. 2023).

Despite significant advances, existing meta-learning BO methods face critical limitations when dealing with heterogeneous task distributions. Most approaches optimize for average performance across tasks, potentially failing on challenging outlier scenarios that may be mission-critical in practice. Recent likelihood-free approaches like MALIBO (Pan et al. 2024) demonstrate improved robustness to task heterogeneity but still lack principled mechanisms for worst-case performance guarantees. Additionally, while diversity sampling techniques have shown promise for individual BO runs (Nava, Mutny, and Krause 2022), they have not been systematically integrated with meta-learning frameworks to address the fundamental challenge of robust knowledge transfer across diverse optimization landscapes.

Robust Optimization and Game-Theoretic Frameworks

Distributional Robust Optimization (DRO) addresses distributional uncertainty by minimizing worst-case expected loss over uncertainty sets (Lin, Fang, and Gao 2022; Ben-Tal et al. 2013; Duchi and Namkoong 2019), and has been successfully applied to meta-learning contexts for optimizing challenging tasks rather than average performance (Wang et al. 2022, 2023; Lv et al. 2024). Game-theoretic approaches model strategic interactions between learning algorithms and adversarial task selectors, with recent theoretical frameworks establishing convergence guarantees for tail task risk minimization (Yin et al. 2020; Lv et al. 2024). Complementing these robust optimization principles, Determinantal Point Processes (DPPs) provide principled frameworks for diversity-aware subset selection (Kulesza, Taskar et al. 2012; Kulesza and Taskar 2011), enabling quality-diversity balanced sampling that has improved exploration in BO contexts (Nava, Mutny, and Krause 2022).

However, three critical gaps limit current approaches: existing DRO applications to BO focus on single-task settings (Kirschner et al. 2020) rather than cross-task knowledge transfer; game-theoretic meta-learning frameworks lack adaptation to sequential decision-making contexts where sample efficiency is paramount; and diversity-aware sampling has not been integrated with robust meta-learning principles for heterogeneous task collections. Our work bridges these gaps by developing the first hierarchical game-theoretic framework that combines CVaR-based robust task selection, diversity-aware sample learning, and uncertainty-aware adaptation, providing principled mechanisms for worst-case performance optimization while maintaining competitive average-case results in meta-learning BO scenarios.

Preliminary

This section establishes the foundational concepts for our hierarchical game-theoretic meta-learning framework. We present three core components that work synergistically: CVaR-based robust task selection for identifying challenging scenarios, weighted classification with diversity sampling for efficient sample learning, and uncertainty-aware adaptation for robust target task generalization. The integration of these components enables principled optimization for worst-case performance while maintaining computational efficiency.

CVaR-Based Robust Task Selection. Traditional meta-learning optimizes average performance across tasks, potentially neglecting challenging scenarios that are critical in practice. To address this limitation, we employ Conditional Value-at-Risk (CVaR) as our primary risk measure for identifying and prioritizing challenging tasks. For a loss variable L, CVaR at confidence level α_{task} captures the expected loss in the worst $(1 - \alpha_{\text{task}})$ fraction of cases:

$$\text{CVaR}_{\alpha_{\text{task}}}[L] = \mathbb{E}[L \mid L \geq \text{VaR}_{\alpha_{\text{task}}}[L]], \tag{1}$$

where $\text{VaR}_{\alpha_{\text{task}}}[L]$ is the Value-at-Risk threshold. The key insight is leveraging CVaR's dual representation to reformulate discrete task selection as a continuous optimization problem. This enables us to identify challenging task subsets through an adaptive threshold ξ that separates high-loss from low-loss tasks, expressed as:

$$\xi^* + \frac{1}{1-\alpha_{\text{task}}}\mathbb{E}_{p(\tau)}[\max\{L(\tau;\theta) - \xi^*, 0\}], \tag{2}$$

where ξ^* automatically adapts to focus meta-learning on the most informative challenging scenarios.

Weighted Classification with Diversity Sampling. Within selected challenging tasks, we must determine which training samples provide the most learning value while maintaining exploration diversity. We adopt a weighted classification framework that initializes sample weights based on improvement potential and evolves them through learning feedback. Sample weights begin with utility-based initialization where promising samples (below a quantile threshold y_{th}) receive weights proportional to their improvement potential:

$$w_{0,i} = \begin{cases} y_{\text{th}} - y_i & \text{if } y_i < y_{\text{th}} \\ \text{constant} & \text{otherwise} \end{cases} \tag{3}$$

To prevent redundant sampling of similar high-value points, we employ k-Determinantal Point Processes (k-DPPs) with diagonal kernel $K_t = \text{diag}(w_t)$, which naturally balances quality (high weights) and diversity (repulsive selection).

Uncertainty-Aware Task Adaptation. A fundamental challenge in meta-learning is adapting to target tasks that may differ significantly from the meta-training distribution. Deterministic representations can lead to overconfident predictions and poor exploration on novel tasks. We address this through probabilistic task embeddings that explicitly model and leverage uncertainty. Task embeddings follow a prior distribution $z_\tau \sim \mathcal{N}(0, I)$, and we perform Bayesian inference to estimate the posterior after observing target task data D_N:

$$p(z_\tau \mid D_N) \approx \mathcal{N}(z_{\text{MAP}}, \Sigma_N), \tag{4}$$

where z_{MAP} represents the most likely task embedding and Σ_N captures uncertainty through the Laplace approximation. This enables Thompson sampling for principled exploration: when task uncertainty is high, the method explores more broadly; as confidence increases, it exploits meta-learned knowledge more aggressively. This uncertainty-aware mechanism ensures robust adaptation even when target tasks exhibit significant distribution shifts from meta-training data.

Methodology

This section presents our hierarchical game-theoretic meta-learning framework for robust Bayesian optimization. We first formalize the problem and establish the hierarchical structure, then detail the task-level robustness mechanism via Stackelberg games, followed by sample-level learning through weighted classification, and conclude with the integrated algorithm and theoretical analysis.

Problem Formulation and Hierarchical Game-Theoretic Framework

We formalize the meta-learning Bayesian optimization problem within a hierarchical framework as illustrated in Figure 2. Let $\mathcal{T}$ denote a distribution over optimization tasks, where each task $\tau \sim \mathcal{T}$ corresponds to optimizing a black-box function $f_\tau : \mathcal{X} \to \mathbb{R}$ over domain $\mathcal{X} \subset \mathbb{R}^d$ (where d is the input dimension). Following the likelihood-free framework, we represent the acquisition function as $\alpha_\tau(x; \theta, z_\tau)$, where $\theta \in \Theta$ denotes meta-parameters and $z_\tau \in \mathbb{R}^{d_{\text{embed}}}$ represents task-specific embeddings (with d_{embed} being the embedding dimension).

To address the limitations of average-case optimization, we propose a hierarchical learning structure operating at two complementary levels. At the task level, we employ a Stackelberg game between a task selector (leader) and meta-learner (follower), where the selector strategically identifies challenging task distributions while the learner develops robust solutions. At the sample level, we implement weighted classification with diversity-aware sampling to emphasize high-value training samples within each selected task. This decomposition enables principled curriculum learning where challenging tasks guide learning direc-

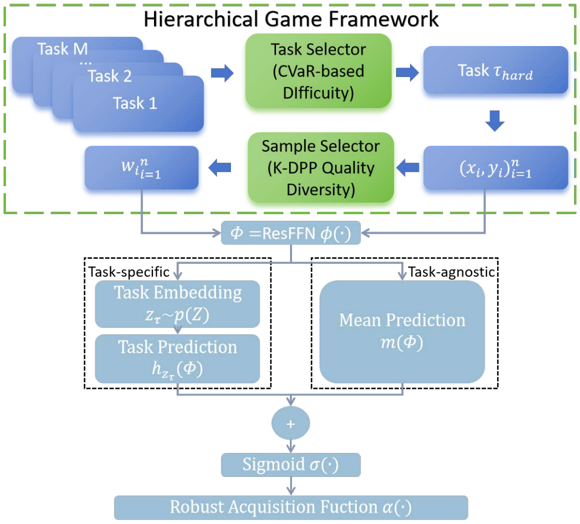

Figure 2: Overview of MetaGameBO's hierarchical game-theoretic framework. The system integrates CVaR-based task selection and k-DPP sample selection to generate robust acquisition functions through task-specific and task-agnostic pathways.

tion and strategic sample selection ensures efficient parameter updates.

The task-level equilibrium is characterized by:

$$q^*_{\text{task}} = \arg \max_{q_{\text{task}} \in \mathcal{Q}_{\alpha_{\text{task}}}} \mathbb{E}_{q_{\text{task}}}[L(\tau; \theta^*(q_{\text{task}}))], \tag{5}$$

where q_{task} represents the task distribution, $\mathcal{Q}_{\alpha_{\text{task}}}$ is the uncertainty set focusing on the $(1-\alpha_{\text{task}})$ fraction of most challenging tasks (with $\alpha_{\text{task}} \in [0,1]$ being the confidence level), $\theta^*(q_{\text{task}}) = \arg\min_\theta \mathbb{E}_{q_{\text{task}}}[L(\tau;\theta)]$ represents the meta-learner's optimal response, and $L(\tau;\theta)$ denotes the loss function for task τ under parameters θ. This hierarchical structure ensures meta-learning focuses on challenging scenarios while maintaining computational tractability.

Task-Level Robustness via Stackelberg Games

A fundamental challenge in meta-learning Bayesian optimization is selecting informative meta-training tasks. Standard approaches either treat all tasks equally or use simple heuristics, leading to suboptimal meta-parameter learning that fails on challenging scenarios or overfits to extreme cases.

To address this systematically, we formulate task selection as a Stackelberg game between a task selector (leader) and meta-learner (follower). The task selector strategically chooses challenging tasks to maximize learning robustness, while the meta-learner optimizes parameters on the selected distribution. This adversarial setup ensures meta-parameters develop robustness against difficult scenarios while maintaining good average-case performance.

The task selector solves Eq. (5), where $\theta^*(q_{\text{task}}) = \arg\min_{\theta\in\Theta} \mathbb{E}_{q_{\text{task}}}[L(\tau;\theta)]$ is the meta-learner's best response. Leveraging the CVaR dual representation introduced in Preliminary, we reformulate this discrete task selection into continuous optimization. Specifically, we employ the dual formulation to identify an adaptive threshold $\xi \in \mathbb{R}$ that automatically separates challenging tasks ($L(\tau;\theta) \geq \xi$) from easier tasks, where ξ corresponds to the α_{task}-quantile of the task loss distribution.

In practice, given M meta-training tasks $\{\tau_i\}_{i=1}^M$ and current meta-parameters θ_t at iteration t, we evaluate losses $\{L_i = L(\tau_i;\theta_t)\}_{i=1}^M$ and estimate the threshold $\hat{\xi}_{\alpha_{\text{task}}}$ using kernel density estimation with Gaussian kernels. The challenging task subset is dynamically updated as $\mathcal{T}^{(t)}_{\text{hard}} = \{\tau_i \mid L_i \geq \hat{\xi}_{\alpha_{\text{task}}}\}$. The Stackelberg equilibrium is computed through alternating optimization: the meta-learner updates parameters via gradient descent, while the task selector updates the distribution by solving CVaR maximization.

Sample-Level Learning via Weighted Classification

Within each selected challenging task, determining how to emphasize different training samples remains critical for effective meta-learning. Standard approaches either treat all samples equally or apply simple heuristics, missing opportunities to leverage varying learning value across samples. The fundamental challenge lies in balancing three objectives: emphasizing high-value samples, maintaining diversity to prevent overfitting, and adapting weights dynamically.

We adopt a weighted classification approach (Tiao et al. 2021; Song et al. 2022) that integrates utility-based initialization, diversity-aware sampling, and adaptive evolution. For each task with observations (x_i, y_i) where $i \in \{1,\ldots,n_\tau\}$ (with n_τ being the number of samples), we establish a utility threshold $y_{\text{th}} = \text{quantile}(y,\gamma)$ where $\gamma \in [0,1]$ is the quantile parameter. Promising samples ($y_i < y_{\text{th}}$) receive weights proportional to improvement potential: $w_{1,i} = (y_{\text{th}} - y_i)$, while reference samples ($y_i \geq y_{\text{th}}$) receive baseline weights. After normalization and class-balancing (where s_1 and s_0 represent counts of promising and reference samples), final weights are obtained as $w \leftarrow w/\text{mean}(w)$.

While utility-based weighting identifies high-value samples, purely weight-based selection can lead to redundant sampling of similar points. To address this, we employ k-Determinantal Point Processes (k-DPP) that simultaneously respect quality preferences and promote diversity. The determinantal structure naturally introduces repulsive interactions between similar items, while the diagonal kernel $K_t = \text{diag}(w_t)$ (where w_t represents the weight vector at iteration t) encodes quality preferences (Kulesza, Taskar et al. 2012; Dereziński 2019). This provides a principled solution balancing quality and diversity with theoretical guarantees.

As meta-learning progresses, sample weights evolve through multiplicative updates reflecting actual learning value:

$$w_{t+1,i} = w_{t,i} \exp\left(\eta_s \frac{\ell_{t,i}(\theta_t) - \bar{\ell}_t}{q_{t,i}}\right), \tag{6}$$

where $\eta_s > 0$ is the sample-level learning rate, $\ell_{t,i}(\theta_t)$ is the loss for sample i under parameters θ_t, $\bar{\ell}_t$ provides normalization, and $q_{t,i}$ is the k-DPP sampling probability for importance sampling correction.

Integrated Algorithm with Uncertainty-Aware Adaptation

The hierarchical components integrate to form MetaGameBO's end-to-end framework that addresses target task adaptation through uncertainty-aware mechanisms. A fundamental challenge is leveraging meta-knowledge while adapting to target tasks that may differ significantly from the meta-training distribution. Standard deterministic representations can lead to overconfident predictions and poor exploration when target tasks exhibit novel characteristics.

Building upon the uncertainty modeling framework in MALIBO (Pan et al. 2024), we develop an enhanced adaptation mechanism that integrates with our hierarchical game-theoretic approach. Our framework extends MALIBO's probabilistic task embeddings through principled robustness mechanisms. We model task embeddings as $z_\tau \sim \mathcal{N}(0, I_{d_{\text{embed}}})$ and employ variational inference during meta-training. The empirical distribution is regularized to match the prior through $\mathcal{L}_{\text{reg}} = \beta \cdot \text{KL}[q_{\text{emp}}(z) \| p(z)]$ where $\beta > 0$ is the regularization weight, preventing overfitting to meta-training tasks.

For target adaptation, we employ Bayesian inference to estimate the posterior $p(z_\tau | D_N)$ after observing N evaluations $D_N = \{(x_j, y_j)\}_{j=1}^{N}$. Using Laplace approximation, we obtain $q(z_\tau) = \mathcal{N}(z_{\text{MAP}}, \Sigma_N)$, where Σ_N captures uncertainty through the inverse Hessian. Thompson sampling leverages this uncertainty by drawing $\hat{z}_\tau \sim q(z_\tau)$ for acquisition functions, providing exploration that adapts to task uncertainty. To ensure robustness against distribution mismatch, we integrate gradient boosting residuals: $\hat{f}(x) = f_{\text{meta}}(x; \theta, z_\tau) + f_{\text{boost}}(x)$, automatically balancing meta-learned and target-specific components.

Algorithm 1 presents our approach, with meta-training alternating between CVaR-based task selection and weighted sample learning, while target adaptation combines Bayesian inference with Thompson sampling. Complete implementation details and algorithm are provided in the Appendix.

Theoretical Analysis

This section provides theoretical guarantees for MetaGameBO's hierarchical game-theoretic framework, establishing convergence properties, generalization bounds, and sample complexity results. Our analysis demonstrates that CVaR-based robust optimization provides superior worst-case guarantees, while uncertainty-aware meta-learning achieves faster convergence than standard approaches.

Our theoretical analysis assumes standard regularity conditions: (1) the loss function $L(\tau; \theta)$ is L_L-Lipschitz continuous and μ-strongly convex in θ; (2) the parameter space Θ and CVaR constraint set $\mathcal{Q}_{\alpha_{\text{task}}}$ are compact; and (3) task embeddings satisfy $\|z_\tau\|_2 \leq R$ for some constant $R > 0$.

The meta-training algorithm seeks a local Stackelberg equilibrium through alternating optimization between task selector and meta-learner. Unlike standard meta-learning

Algorithm 1: MetaGameBO (Simplified)

Input: Meta-training tasks $\{\tau_i\}_{i=1}^{M}$, confidence level α_{task}, learning rates η, η_s
Initialize: Meta-parameters θ_0, sample weights w_0
Meta-training:
for $t = 1$ to T_{meta} **do**
  Evaluate task losses and estimate CVaR threshold $\hat{\xi}_{\alpha_{\text{task}}}$
  Select challenging tasks: $\mathcal{T}_{\text{hard}}^{(t)} \leftarrow \{\tau_i \mid L_i \geq \hat{\xi}_{\alpha_{\text{task}}}\}$
  for each $\tau \in \mathcal{T}_{\text{hard}}^{(t)}$ **do**
    Initialize utility-based weights and apply k-DPP sampling
    Update sample weights using Eq. (6)
  end for
  Update meta-parameters: $\theta_t \leftarrow \theta_{t-1} - \eta \nabla_\theta \mathcal{L}_{\text{weighted}}$
end for
Target adaptation:
for each evaluation step **do**
  Estimate task embedding posterior via Bayesian inference
  Sample $\hat{z}_\tau$ using Thompson sampling
  Generate acquisition function and select next candidate
  Evaluate and update gradient boosting component
end for
Output: Optimized solution for target task

that treats all tasks equally, our framework creates strategic interaction where the task selector chooses challenging tasks while the meta-learner adapts to handle difficult scenarios.

Theorem 1 (Convergence to Local Stackelberg Equilibrium) *Under the regularity conditions stated above, the alternating optimization procedure converges to a local Stackelberg equilibrium* $(\mathcal{T}_{\mathit{hard}}^*, \theta^*)$ *at rate:*

$$\mathbb{E}[\|\theta^{(t)} - \theta^*\|^2] \leq O\left(\frac{\log t}{t}\right). \tag{7}$$

A key question is how meta-learned parameters generalize to unseen target tasks. Our CVaR-based approach provides stronger guarantees through distributional robustness. By focusing on challenging tasks in the upper tail of the loss distribution, we create robust learning that handles distribution shift between meta-training and target tasks.

Theorem 2 (Meta-Learning Generalization Bound) *Let* $\mathcal{T}_{\mathit{test}}$ *be test tasks drawn from the same distribution as meta-training tasks. With probability at least* $1 - \delta$*, the generalization error satisfies:*

$$\begin{aligned}\mathbb{E}_{\tau \sim \mathcal{T}_{\mathit{test}}}[L(\tau; \theta^*)] &\leq \hat{L}_{\mathit{CVaR}}(\mathcal{T}_{\mathit{hard}}; \theta^*) \\ &\quad + O\left(\sqrt{\frac{d_{\mathit{embed}} \log(T_{\mathit{meta}}/\delta)}{T_{\mathit{meta}}}}\right),\end{aligned} \tag{8}$$

where the effective dimension d_{embed} *replaces ambient dimension* d *due to learned task representation.*

Finally, we analyze target task adaptation sample complexity. Meta-learning provides exponential improvements through learned embeddings: instead of learning in full parameter space of dimension d, we learn in embedding space of dimension $d_{\text{embed}} \ll d$. Combined with Thompson sampling and gradient boosting, this leads to significant efficiency gains.

Theorem 3 (Target Task Sample Complexity) *For target task adaptation with MetaGameBO, achieving ϵ-optimal performance with probability $1-\delta$ requires:*

$$N = O\left(\epsilon^{-2}\log(d_{embed}/\delta)\cdot\log(1/\epsilon)\right) \tag{9}$$

samples, compared to $O(\epsilon^{-2}d)$ for standard BO without meta-learning.

These results establish that MetaGameBO enjoys strong theoretical properties while achieving logarithmic improvement in dimension dependence. The CVaR-based training ensures robustness against challenging tasks, while the hierarchical structure maintains computational efficiency. Complete proofs are provided in the Appendix.

Experiments

In this section, we conduct comprehensive experiments to validate the effectiveness of MetaGameBO across multiple dimensions. Our experimental evaluation follows a systematic approach: we first establish the experimental setup and baseline comparisons, then demonstrate the core methodology through controlled experiments on synthetic functions, followed by extensive evaluation on real-world AutoML benchmarks. Finally, we provide detailed ablation studies to analyze the contribution of each component in our hierarchical framework. Due to space limitations, detailed benchmark descriptions, hyperparameter sensitivity analysis, Robustness analysis, and computational complexity with runtime analysis are provided in the Appendix. Code: https://github.com/Allen0497/MetaGameBO.

Baselines and Evaluation Metrics

We compare MetaGameBO against multiple representative baselines spanning different methodological approaches to ensure comprehensive evaluation. For non-meta-learning methods, we include Random Search (Bergstra and Bengio 2012) as a fundamental baseline and Gaussian Process-based BO (GP) (Snoek, Larochelle, and Adams 2012) with Expected Improvement acquisition function as the standard BO approach. We also evaluate LFBO (Song et al. 2022), a likelihood-free method that directly learns acquisition functions.

For meta-learning BO methods, we select established approaches that represent different technical paradigms: ABLR (Perrone et al. 2018) combines neural networks with Bayesian linear regression for scalable adaptation; RGPE (Feurer et al. 2018) uses ranking-weighted GP ensembles to leverage multi-task knowledge; MetaBO (Volpp et al. 2018) employs reinforcement learning to meta-learn acquisition strategies; PFN (Müller et al. 2023) utilizes pre-trained transformers for in-context optimization; and MALIBO (Pan et al. 2024) provides a recent likelihood-free meta-learning approach with uncertainty-aware adaptation.

Method	**Average Loss** ↓	**CVaR** ↓
Random	0.8945 ± 0.00156	1.4532 ± 0.00892
GP	0.7891 ± 0.00134	1.3245 ± 0.00756
LFBO	0.5823 ± 0.00089	0.9876 ± 0.00445
PFN	0.5247 ± 0.00076	0.8934 ± 0.00523
MALIBO	0.4277 ± 0.00048	0.8195 ± 0.00376
MetaGameBO	**0.0183 ± 0.00003**	**0.0934 ± 0.00074**

Table 1: Performance on Forrester function benchmark

Following the evaluation protocol of MALIBO (Pan et al. 2024), we use normalized regret as the primary performance metric, defined as $\min_{x\in X_N}(f^t(x) - f^t_{\min})/(f^t_{\max} - f^t_{\min})$, where X_N represents the set of evaluated inputs up to iteration N, and $f^t_{\min}$, $f^t_{\max}$ are the minimum and maximum values across all available evaluations for task t. This normalization enables fair comparison across tasks with different scales and ranges. Additionally, we report Conditional Value at Risk (CVaR) to assess the tail risk performance, which is particularly relevant for our game-theoretic formulation. All results are averaged over 100 independent runs to ensure statistical reliability, with standard errors reported for significance testing.

Synthetic Benchmarks

We begin our evaluation with the Forrester function (Forrester, Sobester, and Keane 2008), a classic one-dimensional nonlinear test function that exhibits multiple local optima and varying exploration-exploitation trade-offs, making it an ideal testbed for validating our game-theoretic risk-aware framework.

Table 1 presents the performance comparison on the Forrester benchmark. MetaGameBO achieves remarkable performance improvements, reducing average loss by 95.7% compared to MALIBO and 97.9% compared to GP-based optimization. The CVaR results show 88.6% lower tail risk than MALIBO, demonstrating the effectiveness of our risk-aware formulation.

The results clearly reflect methodological advances: non-meta-learning methods (Random, GP) struggle with limited data efficiency, while likelihood-free approaches (LFBO) show modest improvements but lack sophisticated meta-learning mechanisms. Pre-trained methods (PFN) benefit from large-scale pretraining but cannot adapt to specific task distributions. MALIBO represents significant advancement through uncertainty-aware meta-learning, yet the substantial performance gap validates our hypothesis that explicitly optimizing for challenging tasks and informative samples leads to superior meta-learning performance.

Real-world AutoML Benchmarks

We conduct comprehensive evaluation on real-world AutoML optimization tasks to validate the practical effectiveness of MetaGameBO. Our assessment covers three established benchmarks that represent diverse optimization challenges in automated machine learning: NASBench201 (Dong and Yang 2020) for neural architecture search,

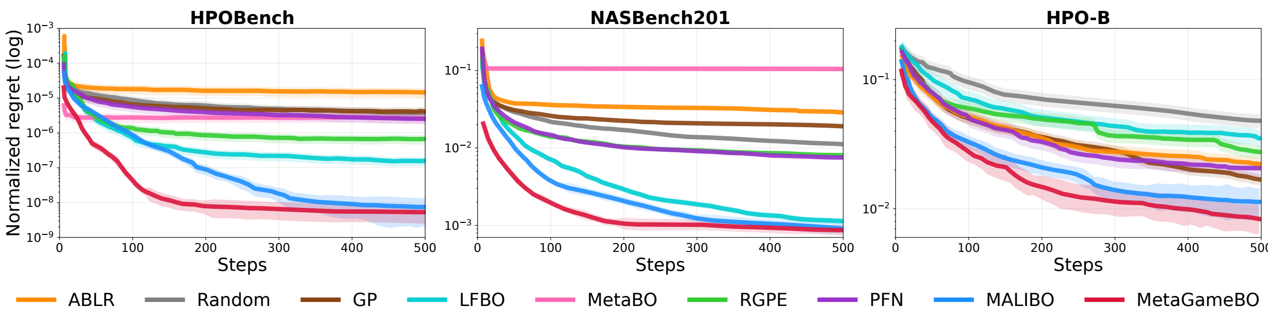

Figure 3: Aggregated normalized regrets for BO algorithms on real-world AutoML problems. MetaGameBO consistently outperforms all baseline methods across HPOBench, NASBench201, and HPO-B benchmarks, demonstrating superior warm-starting capabilities and convergence performance.

HPOBench (Klein and Hutter 2019) for neural network hyperparameter optimization, and HPO-B (Arango et al. 2021) for machine learning algorithm configuration.

NASBench201 involves designing neural cells with 6 discrete parameters across 15,625 unique architectures, evaluated on CIFAR-10, CIFAR-100, and ImageNet-16 datasets with the objective of maximizing validation accuracy. HPOBench focuses on optimizing hyperparameters for two-layer feed-forward regression networks across four UCI datasets, featuring a 9-dimensional continuous search space. HPO-B provides comprehensive evaluation with approximately 6 million hyperparameter evaluations across 16 search spaces of varying dimensionality (2-18 dimensions), targeting accuracy maximization for different machine learning models.

For HPOBench and NASBench201, we employ a leave-one-task-out evaluation strategy where each task serves as the target while others provide meta-training data. We construct meta-datasets by randomly sampling 512 configuration-objective pairs from related tasks to ensure fair comparison across methods. For HPO-B, we utilize the provided meta-train/meta-validation splits for training and evaluate on the designated test tasks, with performance reported after five fixed random initialization points.

Figure 3 presents the aggregated performance across all benchmarks, where MetaGameBO consistently demonstrates superior anytime performance, achieving faster convergence and better final solutions compared to all baselines. Traditional GP-based approaches struggle with the heterogeneous nature of AutoML problems, particularly in HPOBench where regression objectives exhibit abrupt changes that violate standard GP assumptions. Existing meta-learning methods like ABLR, RGPE, and FSBO demonstrate limited effectiveness due to their requirement for extensive meta-data, making them less suitable for scenarios with limited related tasks.

MALIBO delivers strong performance via meta-learning, showing steady gains over baseline methods. However, MetaGameBO substantially outperforms MALIBO across all benchmarks, with improvements in warm-starting efficiency and convergence speed. The superior performance stems from our game-theoretic framework's systematic focus on challenging optimization scenarios while ensuring diverse and informative sample selection, enabling more effective knowledge transfer in AutoML contexts where task diversity and limited evaluation budgets demand efficient optimization strategies.

Ablation Study

To understand the individual contributions of our hierarchical framework components, we conduct systematic ablation studies on the Forrester benchmark. We evaluate three key variants: the original MALIBO baseline, MetaGameBO with only task-level selection (CVaR-based challenging task identification), MetaGameBO with only sample-level selection (k-DPP weighted sampling), and the complete MetaGameBO framework. Task-level selection alone achieves meaningful improvements, reducing average loss by 21.7% and CVaR by 28.1%. Sample-level selection shows dramatically stronger impact, achieving 71.8% reduction in average loss and 70.5% reduction in CVaR, highlighting the critical importance of intelligent sample weighting and diversity-aware selection. The complete MetaGameBO framework achieves the best performance, with sample-level selection being the primary driver while task-level mechanisms provide complementary benefits. This substantial performance gap suggests that strategic sample selection fundamentally transforms meta-learning effectiveness by maximizing information extraction from each training sample. Detailed results are presented in the appendix.

Conclusion

This paper introduced MetaGameBO, a hierarchical game-theoretic meta-learning framework that formulates meta-learning as CVaR-based task selection through Stackelberg games and weighted classification with k-DPP sampling for robust Bayesian optimization. We established theoretical guarantees including convergence to local Stackelberg equilibria and improved sample complexity compared to standard BO methods. Extensive experiments demonstrated substantial performance improvements over state-of-the-art baselines across both synthetic benchmarks and real-world AutoML tasks.

Acknowledgements

This work was partly supported by The National Key Research and Development Program of China (2024YFE0202900); Beijing Natural Science Foundation (4244096); The National Natural Science Foundation of China under Grant (62436001,62406019, 62536001, 62176020); The Joint Foundation of the Ministry of Education for Innovation team (8091B042235); The State Key Laboratory of Rail Traffic Control and Safety (RCS2023K006); the Talent Fund of Beijing Jiaotong University (2024XKRC075); Natural Science Foundation of Chongqing (Grant CSTB2023NSCQ-MSX1020).

References

Arango, S. P.; Jomaa, H. S.; Wistuba, M.; and Grabocka, J. 2021. HPO-B: A Large-Scale Reproducible Benchmark for Black-Box HPO based on OpenML. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*.

Bardenet, R.; Brendel, M.; Kégl, B.; and Sebag, M. 2013. Collaborative hyperparameter tuning. In *International conference on machine learning*, 199–207. PMLR.

Ben-Tal, A.; Den Hertog, D.; De Waegenaere, A.; Melenberg, B.; and Rennen, G. 2013. Robust solutions of optimization problems affected by uncertain probabilities. *Management Science*, 59(2): 341–357.

Bergstra, J.; Bardenet, R.; Bengio, Y.; and Kégl, B. 2011. Algorithms for hyper-parameter optimization. *Advances in neural information processing systems*, 24.

Bergstra, J.; and Bengio, Y. 2012. Random search for hyper-parameter optimization. *The journal of machine learning research*, 13(1): 281–305.

Brochu, E.; Cora, V. M.; and De Freitas, N. 2010. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*.

Dereziński, M. 2019. Fast determinantal point processes via distortion-free intermediate sampling. In *Conference on Learning Theory*, 1029–1049. PMLR.

Dong, X.; and Yang, Y. 2020. NAS-Bench-201: Extending the Scope of Reproducible Neural Architecture Search. In *International Conference on Learning Representations*.

Duchi, J.; and Namkoong, H. 2019. Variance-based regularization with convex objectives. *Journal of Machine Learning Research*, 20(68): 1–55.

Feurer, M.; Letham, B.; and Bakshy, E. 2018. Scalable meta-learning for bayesian optimization using ranking-weighted gaussian process ensembles. In *AutoML Workshop at ICML*, volume 7, 5.

Feurer, M.; Letham, B.; Hutter, F.; and Bakshy, E. 2018. Practical transfer learning for bayesian optimization. *arXiv preprint arXiv:1802.02219*.

Feurer, M.; Springenberg, J.; and Hutter, F. 2015. Initializing bayesian hyperparameter optimization via meta-learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 29.

Finn, C.; Rajeswaran, A.; Kakade, S.; and Levine, S. 2019. Online meta-learning. In *International conference on machine learning*, 1920–1930. PMLR.

Finn, C.; Xu, K.; and Levine, S. 2018. Probabilistic model-agnostic meta-learning. *Advances in neural information processing systems*, 31.

Forrester, A.; Sobester, A.; and Keane, A. 2008. *Engineering design via surrogate modelling: a practical guide*. John Wiley & Sons.

Frazier, P. I.; and Wang, J. 2016. Bayesian optimization for materials design. *Information science for materials discovery and design*, 45–75.

Griffiths, R.-R.; and Hernández-Lobato, J. M. 2020. Constrained Bayesian optimization for automatic chemical design using variational autoencoders. *Chemical science*, 11(2): 577–586.

Hospedales, T.; Antoniou, A.; Micaelli, P.; and Storkey, A. 2021. Meta-learning in neural networks: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 44(9): 5149–5169.

Jamal, M. A.; and Qi, G.-J. 2019. Task agnostic meta-learning for few-shot learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 11719–11727.

Jin, H.; Song, Q.; and Hu, X. 2019. Auto-keras: An efficient neural architecture search system. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, 1946–1956.

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

Kandasamy, K.; Neiswanger, W.; Schneider, J.; Poczos, B.; and Xing, E. P. 2018. Neural architecture search with bayesian optimisation and optimal transport. *Advances in neural information processing systems*, 31.

Kirschner, J.; Bogunovic, I.; Jegelka, S.; and Krause, A. 2020. Distributionally robust Bayesian optimization. In *International Conference on Artificial Intelligence and Statistics*, 2174–2184. PMLR.

Klein, A.; and Hutter, F. 2019. Tabular benchmarks for joint architecture and hyperparameter optimization. *arXiv preprint arXiv:1905.04970*.

Kulesza, A.; and Taskar, B. 2011. k-dpps: Fixed-size determinantal point processes. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, 1193–1200.

Kulesza, A.; Taskar, B.; et al. 2012. Determinantal point processes for machine learning. *Foundations and Trends® in Machine Learning*, 5(2–3): 123–286.

Lin, F.; Fang, X.; and Gao, Z. 2022. Distributionally robust optimization: A review on theory and applications. *Numerical Algebra, Control and Optimization*, 12(1): 159–212.

Lv, Y.; Wang, C.; Liang, D.; and Xie, Z. 2024. Theoretical Investigations and Practical Enhancements on Tail Task Risk Minimization in Meta Learning. In *The Thirty-eighth*

Annual Conference on Neural Information Processing Systems.

Müller, S.; Feurer, M.; Hollmann, N.; and Hutter, F. 2023. Pfns4bo: In-context learning for bayesian optimization. In *International Conference on Machine Learning*, 25444–25470. PMLR.

Nava, E.; Mutny, M.; and Krause, A. 2022. Diversified sampling for batched Bayesian optimization with determinantal point processes. In *International Conference on Artificial Intelligence and Statistics*, 7031–7054. PMLR.

Pan, J.; Falkner, S.; Berkenkamp, F.; and Vanschoren, J. 2024. MALIBO: meta-learning for likelihood-free bayesian optimization. In *Proceedings of the 41st International Conference on Machine Learning*, 39102–39134.

Perrone, V.; Jenatton, R.; Seeger, M. W.; and Archambeau, C. 2018. Scalable hyperparameter transfer learning. *Advances in neural information processing systems*, 31.

Rothfuss, J.; Koenig, C.; Rupenyan, A.; and Krause, A. 2023. Meta-learning priors for safe Bayesian optimization. In *Conference on robot learning*, 237–265. PMLR.

Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R. P.; and De Freitas, N. 2015. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1): 148–175.

Snoek, J.; Larochelle, H.; and Adams, R. P. 2012. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, 25.

Song, J.; Yu, L.; Neiswanger, W.; and Ermon, S. 2022. A general recipe for likelihood-free Bayesian optimization. In *International conference on machine learning*, 20384–20404. PMLR.

Srinivas, N.; Krause, A.; Kakade, S.; and Seeger, M. 2010. Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design. In *Proceedings of the 27th International Conference on Machine Learning*, 1015–1022. Omnipress.

Tiao, L. C.; Klein, A.; Seeger, M. W.; Bonilla, E. V.; Archambeau, C.; and Ramos, F. 2021. BORE: Bayesian optimization by density-ratio estimation. In *International conference on machine learning*, 10289–10300. PMLR.

Vanschoren, J. 2018. Meta-learning: A survey. *arXiv preprint arXiv:1810.03548*.

Volpp, M.; Fröhlich, L. P.; Fischer, K.; Doerr, A.; Falkner, S.; Hutter, F.; and Daniel, C. 2018. Meta-Learning Acquisition Functions for Transfer Learning in Bayesian Optimization. In *International Conference on Learning Representations*.

Wang, Q.; Lv, Y.; Xie, Z.; Huang, J.; et al. 2023. A simple yet effective strategy to robustify the meta learning paradigm. *Advances in Neural Information Processing Systems*, 36: 12897–12928.

Wang, Y.; and Arora, R. 2024. On the stability and generalization of meta-learning. *Advances in Neural Information Processing Systems*, 37: 83665–83710.

Wang, Z.; Wang, X.; Shen, L.; Suo, Q.; Song, K.; Yu, D.; Shen, Y.; and Gao, M. 2022. Meta-learning without data via wasserstein distributionally-robust model fusion. In *Uncertainty in Artificial Intelligence*, 2045–2055. PMLR.

Wistuba, M.; Schilling, N.; and Schmidt-Thieme, L. 2018. Scalable gaussian process-based transfer surrogates for hyperparameter optimization. *Machine Learning*, 107(1): 43–78.

Yin, M.; Tucker, G.; Zhou, M.; Levine, S.; and Finn, C. 2020. Meta-Learning without Memorization. In *International Conference on Learning Representations*.

Yu, T.; and Zhu, H. 2020. Hyper-parameter optimization: A review of algorithms and applications. *arXiv preprint arXiv:2003.05689*.

Zhou, H.; Ma, X.; and Blaschko, M. B. 2024. A corrected expected improvement acquisition function under noisy observations. In *Asian Conference on Machine Learning*, 1747–1762. PMLR.