Proceedings of the ASME 2026
International Design Engineering Technical Conferences and
Computers and Information in Engineering Conference
IDETC/CIE2026
August 24–26, 2026, Houston, TX

IDETC/CIE 2026-193982

ROBUST COMPETITIVE BAYESIAN OPTIMIZATION

Siyu Chen[1], Zhenghui Sha[1,*]

[1]Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX

ABSTRACT

Sequential design decisions in engineering often arise under both strategic competition and model uncertainty, where agents must optimize performance relative to evolving benchmarks while operating under incomplete and noisy observations. While competitive Bayesian optimization (CBO) incorporates opponent-aware benchmarking into acquisition decisions, existing approaches remain risk-neutral and can be brittle when posterior predictive uncertainty is affected by limited data, noisy observations, surrogate-model miscalibration, and/or evolving competitive benchmark. This paper proposes Robust Competitive Bayesian Optimization (RCBO), a principled extension of CBO that integrates distributional robustness into competition-aware information acquisition. RCBO evaluates competitive advantage under a Kullback–Leibler (KL) distortion of the Gaussian process posterior, derived via the Donsker–Varadhan variational identity, yielding a robust acquisition function that smoothly interpolates between expected advantage and worst-case evaluation through a tunable risk-sensitivity parameter. To ensure consistency across iterations and varying payoff scales, we further introduce an adaptive risk-scaling mechanism that adjusts robustness based on both posterior uncertainty and relative competitive standing, enabling dynamic balancing of exploration and exploitation. Empirical results on multimodal benchmark functions demonstrate that RCBO consistently improves competitive outcomes, convergence speed, and decision stability compared to risk-neutral baselines, particularly when observations are noisy and the competitive benchmark induces uncertainty in the relative payoff. By unifying competition-awareness with principled ambiguity aversion, RCBO advances Bayesian optimization toward more reliable, stable, and robust decision-making in adversarial and uncertain engineering environments.

Keywords: Robust Bayesian optimization, Competitive optimization, Multi-agent decision-making, Design space exploration

*Corresponding author: zsha@austin.utexas.edu
Documentation for `asmeconf.cls`: Version 1.46, May 18, 2026.

1. INTRODUCTION

Real-world sequential decision-making is rarely performed in isolation. Engineering teams and autonomous agents must often decide *what to evaluate next* under limited budgets, noisy feedback, and incomplete information, while simultaneously being judged against external benchmarks—frequently an explicit competitor. In such settings, rational decision-making must be robust: it should not only seek high expected performance, but also guard against posterior predictive ambiguity arising from limited observations, observation noise, possible surrogate-model miscalibration, and uncertainties in the evolving competitive benchmark. In decision-making under uncertainty, *robustness* refers to the ability of a decision rule to maintain satisfactory performance under model misspecification or distributional ambiguity. Formally, robust optimization evaluates a decision by its worst-case performance over an uncertainty set of plausible models or probability distributions [1, 2]. In this sense, robustness provides protection against adverse realizations that may arise when the nominal model is inaccurate. In this paper, we use robustness in this relatively narrower posterior-predictive sense. That is, RCBO protects acquisition decisions against local distortions of the GP-induced payoff distribution, rather than claiming protection against arbitrary structural changes in the physical objective or environment.

A large body of work in robust control and decision theory formalizes this principle by evaluating actions under adversarial perturbations of a nominal belief. Early formulations include minimax decision theory [3] and maxmin expected utility [4], while modern treatments characterize ambiguity aversion via divergence-penalized or variational preferences [5, 6]. In engineering optimization, distributionally robust optimization (DRO) evaluates decisions against worst-case distributions within divergence-based uncertainty sets, including KL balls [7, 8]. KL-based formulations are particularly attractive because they admit a tractable entropic-risk representation, which evaluates a decision by combining its expected payoff with aversion to adverse outcomes. Through exponential tilting, this representation reweights

the nominal distribution toward unfavorable scenarios, enabling a principled trade-off between performance and conservatism.

Bayesian optimization (BO) is a widely used framework for sample-efficient optimization of expensive black-box functions, leveraging Gaussian process surrogates and acquisition functions to balance exploration and exploitation [9–12]. However, standard BO is intrinsically *single-agent* and *self-referential*: improvement is measured relative to an agent's own incumbent best value, implicitly assuming that performance is absolute and exogenous. This formulation overlooks settings in which performance is inherently *relative*—for example, when design quality is evaluated against an evolving competitor or benchmark. In such cases, the payoff functional becomes benchmark-dependent and endogenous, and classical BO provides no normative mechanism for adapting sampling decisions to the competitive state.

To address this limitation, recent work has begun to model competitive sequential decision-making using multi-agent and game-theoretic perspectives, including dueling bandits [13, 14], stochastic games [15], and multi-agent reinforcement learning [16]. Motivated by costly-evaluation settings, our prior work introduced Competitive Bayesian Optimization (CBO) [17], which replaces self-referential improvement with an opponent-aware benchmark and selects evaluations by maximizing *Expected Advantage* (EA). By coupling sampling decisions to the current performance gap, CBO provides a principled answer to *where to sample under competition*.

Despite these advances, a critical gap remains: *robustness under competitive uncertainty*. In competitive Bayesian optimization, uncertainty arises not only from the surrogate model used to estimate the objective but also from the evolving competitive state against which performance is evaluated. These two sources of uncertainty jointly influence the reliability of acquisition decisions and motivate the need for robustness. We therefore distinguish two key challenges in competitive BO. *First, model uncertainty.* As in standard BO, decisions depend on a GP posterior that may suffer from surrogate-model miscalibration due to kernel choice, hyperparameter estimation error, and limited data. The sensitivity of acquisition rules to posterior miscalibration is well documented [11, 12, 18], and robust BO variants such as StableOpt [19] and distributionally robust BO [20] mitigate this effect for absolute objectives. However, in a competitive setting, posterior errors distort not only improvement estimates but also the competitive payoff itself, amplifying the impact of surrogate-model miscalibration.

Second, strategic competition. In competitive optimization, decisions are evaluated relative to an evolving benchmark determined by the opponent's performance [15, 16]. The acquisition objective therefore depends not only on the surrogate model but also on the current competitive standing, such as whether an agent is leading or trailing. Existing competitive BO approaches, including CBO, treat acquisition decisions as risk-neutral with respect to this relative payoff. As a result, the algorithm does not adjust its risk attitude based on the competitive state—for example, adopting more conservative decisions to protect an incumbent lead or more aggressive exploration when falling behind. These observations motivate the following question: *How should Bayesian optimization adapt its sampling decisions when both model uncertainty and competitive standing influence the acquisition process?* In particular, how can robustness be incorporated into competition-aware acquisition functions to stabilize decisions while preserving sample efficiency?

In this paper, We propose RCBO, a principled extension of CBO, a principled extension of CBO that integrates KL-divergence–based ambiguity aversion directly into competition-aware information acquisition. To account for posterior ambiguity in the GP-induced competitive payoff, we introduce KL-divergence–based ambiguity aversion. Intuitively, ambiguity aversion reflects a decision-maker's preference to guard against the possibility that the assumed probability model may be inaccurate. Instead of trusting the Gaussian-process posterior exactly, we allow for nearby alternative probability distributions that slightly distort this posterior. The Kullback–Leibler (KL) divergence provides a natural way to quantify how far such alternative distributions deviate from the nominal belief. By restricting these alternatives to lie within a KL neighborhood of the posterior, the algorithm evaluates decisions under pessimistic but plausible payoff distributions, thereby reducing sensitivity to surrogate-model miscalibration and unfavorable competitive outcomes.

Instead of optimizing the posterior expectation of competitive advantage, RCBO evaluates advantage under a KL-distorted posterior, yielding an entropic risk functional via the Donsker–Varadhan variational identity [5, 21]. This formulation allows an adversary to pessimistically reweight posterior outcomes with a KL penalty, producing a robust competitive acquisition function that smoothly interpolates between risk-neutral Expected Advantage and worst-case evaluation through a tunable risk-sensitivity parameter. By incorporating divergence-based robustness directly into the evaluation of competitive advantage relative to the benchmark, RCBO allows sampling decisions to account for both model uncertainty and the agent's current competitive standing. By adapting risk sensitivity to the competitive state, RCBO unifies uncertainty-aware robustness and competition-aware decision-making within a single acquisition framework.

The remainder of this paper is organized as follows. Section 2 reviews Gaussian process surrogates and the competitive Bayesian optimization framework. Section 3 develops the proposed Robust Competitive Bayesian Optimization (RCBO) method, including the KL-divergence–based robust evaluation and the resulting competition-aware acquisition function. Section 4 presents the experimental setup and empirical results on benchmark problems. Section 5 discusses the implications and limitations of the proposed approach. Finally, Section 6 concludes the paper and outlines directions for future work.

2. PRELIMINARIES

2.1. Gaussian Process Surrogate

We consider a design space $\mathcal{X} \subset \mathbb{R}^d$ and a black–box objective function $f_i : \mathcal{X} \to \mathbb{R}$ for agent i. We model f_i with a Gaussian process (GP) prior $f_i \sim \mathcal{GP}(m_0, k_0)$ [22], where $m_0 : \mathcal{X} \to \mathbb{R}$ is the prior mean function and $k_0 : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is the positive–definite covariance (kernel) function.

At each iteration t, agent i selects a design $x_{it} \in \mathcal{X}$ and

observes a noisy outcome

$$y_{it} = f_i(x_{it}) + \varepsilon_{it}, \quad \varepsilon_{it} \sim \mathcal{N}(0, v^2), \tag{1}$$

where ε_{it} are i.i.d. Gaussian noises with variance v^2. Thus, v^2 represents an additive homoscedastic observation-noise variance, or nugget term, that is added to the diagonal of the GP covariance matrix. In this study, we do not model multiplicative observation uncertainty or input-dependent noise. After t evaluations, the agent has collected a dataset $\mathcal{D}_{i,t} = \{(x_{is}, y_{is})\}_{s=1}^{t}$.

To compute the posterior, we define the design history $X_{it} = [x_{i1}, \ldots, x_{it}]$, the observation vector $\mathbf{y}_{it} = [y_{i1}, \ldots, y_{it}]^\top$, the kernel matrix $K_{it} = [k_0(x_{ia}, x_{ib})]_{1\le a,b\le t}$, and the covariance vector $\mathbf{k}_{it}(x) = [k_0(x_{i1}, x), \ldots, k_0(x_{it}, x)]^\top$. Conditioned on $\mathcal{D}_{i,t}$, the posterior distribution at a candidate x is $f_i(x) \mid \mathcal{D}_{i,t} \sim \mathcal{N}(\mu_{it}(x), \sigma^2_{i,t}(x))$, determined by GP regression as:

$$\mu_{it}(x) = m_0(x) + \mathbf{k}_{it}(x)^\top \big(K_{it} + v^2 I\big)^{-1} \big(\mathbf{y}_{it} - m_0(X_{it})\big), \tag{2}$$

$$\sigma^2_{i,t}(x) = k_0(x, x) - \mathbf{k}_{it}(x)^\top \big(K_{it} + v^2 I\big)^{-1} \mathbf{k}_{it}(x). \tag{3}$$

The posterior variance is denoted by $\sigma^2_{i,t}(x)$ and the corresponding standard deviation by $\sigma_{i,t}(x)$.

2.2. Competitive Bayesian Optimization

Following our prior study [17], we focus on a competitive setting in which the quality of a design is assessed *relative* to an external performance state. This state may correspond to an explicit opponent (agent $-i$) or to a benchmark process summarizing the best designs identified so far. We denote by f_t^{EA} a scalar *competitive benchmark* at iteration t (e.g., an opponent best-so-far value $f^*_{-i,t}$). For the detailed definition of f_t^{EA}, please check the method from [17]. A brief overview of the CBO framework is provided in Appendix A.

For each candidate $x \in \mathcal{X}$, we define the competition-aware payoff random variable

$$Z_i^{\mathrm{comp}}(x) \;:=\; u\big(f_i(x),\, f_t^{\mathrm{EA}}\big), \tag{4}$$

Here, $u(\cdot, \cdot)$ is a deterministic payoff transformation. At iteration t, the benchmark f_t^{EA} is treated as fixed when evaluating a candidate x, while $f_i(x)$ remains random under the GP posterior. Therefore, $Z_i^{\mathrm{comp}}(x)$ is a posterior-induced random competitive payoff.

We assume that u is monotone in its first argument, so that a larger objective value $f_i(x)$ produces no smaller competitive payoff for a fixed benchmark. A canonical choice recovering the usual Expected Advantage (EA) form is the rectified difference [17]:

$$Z_i^{\mathrm{comp}}(x) = \big(f_i(x) - f_t^{\mathrm{EA}}\big)^+. \tag{5}$$

This distinction is important because uncertainty in $f_i(x)$ propagates through the benchmark-dependent payoff $Z_i^{\mathrm{comp}}(x)$, so a design with high nominal objective value may still have uncertain or unfavorable competitive advantage relative to the current benchmark.

3. METHOD

This section introduces a robust competitive Bayesian optimization that augments the EA framework with ambiguity aversion. We first define a KL-distortion robust evaluation that allows an adversary to pessimistically reweight the GP posterior within a KL penalty, yielding a closed-form entropic risk functional. We then instantiate this evaluation for the competition-aware payoff $Z_i^{\mathrm{comp}}(x)$ to obtain the robust competitive acquisition $\alpha_i^{\mathrm{REA}}(x)$, which is maximized to select the next query. Finally, we summarize the resulting BO loop and practical implementation details.

3.1. Robust Evaluation via KL Distortion

Let μ_t denote the baseline posterior probability measure induced by the GP posterior over the latent objective $f_i(x)$ at iteration t. This distribution represents the model's current belief about the uncertain outcome of evaluating design x. Throughout this paper, *ambiguity* refers to uncertainty about the posterior predictive law used to evaluate the acquisition function. For a fixed candidate x and benchmark f_t^{EA}, the GP posterior induces a distribution over the competitive payoff $Z_i^{\mathrm{comp}}(x)$. RCBO models ambiguity by considering alternative payoff distributions that are absolutely continuous with respect to the nominal GP-induced distribution and remain close to it in KL divergence. Thus, ambiguity is distinct from the additive observation noise in Eq. (1) and from a structural distributional shift of the true objective function. For a candidate x, let $Z(x)$ be the induced random utility under μ_t (we ultimately maximize Z).

To introduce ambiguity aversion, we adopt the standard decision-theoretic perspective that the posterior μ_t may itself be imperfect due to model misspecification, limited data, or noise. Instead of assuming that μ_t is exactly correct, we consider nearby alternative probability distributions that represent plausible perturbations of the nominal belief. For example, if the GP posterior predicts a high payoff with moderate uncertainty, a pessimistic decision-maker may wish to evaluate the design under a slightly distorted distribution that assigns greater probability to unfavorable outcomes. This captures the notion of ambiguity aversion: decisions are evaluated not only under the nominal belief but also under nearby distributions that represent possible model errors.

Formally, we allow an adversary ("nature") to replace the nominal posterior μ_t with an alternative probability distribution ν. The distribution ν represents a distorted belief obtained by reweighting the outcomes under the nominal posterior μ_t, thereby capturing pessimistic scenarios in which unfavorable outcomes receive greater probability mass. The deviation of ν from the nominal belief μ_t is quantified using the Kullback–Leibler (KL) divergence, which measures how far the distorted distribution departs from the posterior belief.

$$D_{\mathrm{KL}}(\nu \| \mu_t) = \mathbb{E}_\nu \left[\log \frac{d\nu}{d\mu_t} \right]. \tag{6}$$

Given an ambiguity budget $\rho \ge 0$, we evaluate $Z(x)$ under the worst-case distribution ν chosen from a neighborhood of the nominal posterior μ_t defined by a KL-divergence constraint. Intuitively, this neighborhood—often called a *KL ball*—contains all probability distributions that differ from the nominal posterior μ_t by at most ρ in KL divergence. In other words, we allow nature

to slightly distort the posterior belief, but only within a bounded level of deviation measured by the KL divergence.

$$\mathcal{R}_\rho(Z(x)) := \inf_{\nu \ll \mu_t :\, D_{\mathrm{KL}}(\nu\|\mu_t) \le \rho} \mathbb{E}_\nu[Z(x)]. \tag{7}$$

When $\rho = 0$, feasibility forces $\nu = \mu_t$, hence $\mathcal{R}_0(Z(x)) = \mathbb{E}_{\mu_t}[Z(x)]$. Increasing ρ enlarges the ambiguity set and yields a more conservative (pessimistic) evaluation.

Define the constraint function $c(\nu) := D_{\mathrm{KL}}(\nu\|\mu_t) - \rho \le 0$ and form the Lagrangian [2]

$$\mathcal{L}(\nu, \lambda) := \mathbb{E}_\nu[Z(x)] + \lambda\big(D_{\mathrm{KL}}(\nu\|\mu_t) - \rho\big), \qquad \lambda \ge 0, \tag{8}$$

where λ is the Lagrange multiplier (dual variable) associated with the KL-divergence constraint. In the multiplier formulation, λ controls how costly it is for an adversary to distort the nominal posterior distribution μ_t. A large value of λ strongly penalizes deviations from μ_t, forcing the distorted distribution ν to remain close to the posterior and producing behavior close to the risk-neutral evaluation. In contrast, a small value of λ allows greater distortion of μ_t, permitting the adversary to place more weight on unfavorable outcomes and yielding a more pessimistic (robust) evaluation.

The constrained formulation in Eq. (7), parameterized by the KL radius ρ, and the multiplier formulation in Eq. (8) parameterized by λ are dual representations of the same robustness concept. Increasing ρ enlarges the KL ball and therefore allows more adversarial distortions of the posterior. In the dual formulation, this corresponds to decreasing λ, which reduces the penalty on KL distortion and leads to a more conservative (robust) evaluation. In other words, reducing λ allows stronger adversarial distortions of the posterior, leading to a more pessimistic evaluation of the payoff and therefore a more robust decision rule.

For each fixed $\lambda \ge 0$, we define the *dual function* as the minimum value of the Lagrangian over all admissible distorted distributions ν. Introducing this dual function allows the constrained robust problem in Eq. (7) to be rewritten in a penalized form that depends only on λ. This step is standard in Lagrangian duality and leads directly to the KL-penalized robust functional used in the following derivation.

$$g(\lambda) := \inf_{\nu \ll \mu_t} \mathcal{L}(\nu, \lambda). \tag{9}$$

$g(\lambda)$ is a lower bound on the primal value for any $\lambda \ge 0$ (weak duality), so the tightest bound is obtained by maximizing $g(\lambda)$.

Expanding (9) using (8) yields

$$g(\lambda) = \underbrace{\inf_{\nu \ll \mu_t} \Big\{\mathbb{E}_\nu[Z(x)] + \lambda D_{\mathrm{KL}}(\nu\|\mu_t)\Big\}}_{:= \mathcal{R}_\lambda(Z(x))} - \lambda\rho, \tag{10}$$

which motivates the KL-penalized (multiplier–robust) functional

$$\mathcal{R}_\lambda(Z(x)) := \inf_{\nu \ll \mu_t} \Big\{\mathbb{E}_\nu[Z(x)] + \lambda\, D_{\mathrm{KL}}(\nu\|\mu_t)\Big\}, \qquad \lambda > 0. \tag{11}$$

A standard variational identity (Donsker–Varadhan; see [5, 23, 24]) yields, for $\lambda > 0$. A detailed derivation is provided in Appendix B.

$$\mathcal{R}_\lambda(Z(x)) = -\lambda \log \mathbb{E}_{\mu_t}\left[\exp\big(-Z(x)/\lambda\big)\right], \tag{12}$$

and the minimizing distribution in (11) is the exponentially tilted measure $\nu^\star_\lambda \ll \mu_t$:

$$\frac{d\nu^\star_\lambda}{d\mu_t}(\omega) = \frac{\exp\big(-Z(x,\omega)/\lambda\big)}{\mathbb{E}_{\mu_t}\left[\exp\big(-Z(x)/\lambda\big)\right]}. \tag{13}$$

Equations (12)–(13) provide a practical interpretation of the robust evaluation. Instead of trusting the posterior distribution μ_t exactly, we allow a pessimistic "adversary" to slightly shift probability mass toward unfavorable outcomes. In other words, outcomes that produce smaller competitive payoff are treated as more likely than the nominal GP posterior suggests. The parameter λ controls how strong this pessimism is. When λ is large, changing the posterior is heavily penalized, so the evaluation remains close to the usual risk-neutral expectation. When λ is small, the adversary is allowed to emphasize low-payoff outcomes more strongly, producing a more conservative (risk-averse) evaluation. From this perspective, $\mathcal{R}_\lambda$ acts as a smooth interpolation between the nominal expected payoff and a worst-case evaluation, providing a tunable level of robustness against model uncertainty.

3.2. Robust Competitive Acquisition Function

In the competitive setting, we take the random competitive payoff to be the competitive advantage $Z(x) = Z_i^{\mathrm{comp}}(x)$. We adopt the multiplier–robust (KL-penalized) evaluation from Section 3.1, treating $\lambda > 0$ as a tunable risk-sensitivity parameter. The resulting *robust competitive acquisition* for agent i at iteration t is

$$\alpha_i^{\mathrm{REA}}(x) := \mathcal{R}_\lambda\big(Z_i^{\mathrm{comp}}(x)\big) = \inf_{\nu \ll \mu_t} \Big\{\mathbb{E}_\nu\big[Z_i^{\mathrm{comp}}(x)\big] + \lambda\, D_{\mathrm{KL}}(\nu\|\mu_t)\Big\}. \tag{14}$$

The inner infimum in (14) represents an adversary that perturbs the posterior distribution μ_t toward unfavorable outcomes, subject to a KL penalty that limits the magnitude of the distortion. The risk parameter λ controls this trade-off: larger λ penalizes deviations more strongly and yields behavior close to the risk-neutral evaluation, whereas smaller λ allows stronger pessimistic reweighting of outcomes and produces a more conservative (robust) evaluation.

Using the variational identity in (12), the infimum in (14) admits the closed form

$$\alpha_i^{\mathrm{REA}}(x) = -\lambda \log \mathbb{E}_{\mu_t}\left[\exp\big(-Z_i^{\mathrm{comp}}(x)/\lambda\big)\right]. \tag{15}$$

In our implementation of competitive BO [17], we set $Z_i^{\mathrm{comp}}(x) = \big(f_i(x) - f_t^{\mathrm{EA}}\big)^+$, yielding

$$\alpha_i^{\mathrm{REA}}(x) = -\lambda \log \mathbb{E}_{\mu_t}\left[\exp\big(-\big(f_i(x) - f_t^{\mathrm{EA}}\big)^+/\lambda\big)\right]. \tag{16}$$

Thus, $\alpha_i^{\mathrm{REA}}(x)$ is an entropic risk adjustment of the posterior distribution of competitive payoff.

At each iteration, agent i selects the next candidate by maximizing the robust acquisition over a finite candidate set. Let $\mathcal{X}_{\mathrm{cand},t} \subset \mathcal{X}$ denote the candidate set available for acquisition maximization at iteration t. In the benchmark experiments, this set is constructed in the normalized design domain, and the same

candidate set is used for RCBO and the baseline method. The next query is selected as

$$x_{i,t+1} \in \arg\max_{x \in \mathcal{X}_{\text{cand},t}} \alpha_i^{\text{REA}}(x). \tag{17}$$

The risk-sensitivity parameter λ controls how costly it is for nature to distort the posterior. As $\lambda \to \infty$, KL distortion is heavily penalized and the robust acquisition reduces to the risk-neutral competitive expected advantage:

$$\lim_{\lambda \to \infty} \alpha_i^{\text{REA}}(x) = \mathbb{E}_{\mu_t}\left[Z_i^{\text{comp}}(x)\right]. \tag{18}$$

Conversely, as $\lambda \downarrow 0$, the log–sum–exp term in (15) concentrates on the smallest realizations of $Z_i^{\text{comp}}(x)$, so α_i^{REA} approaches a worst-case (highly pessimistic) evaluation of the competitive payoff.

Figure 1 summarizes the RCBO overall framework. At iteration t, the observed state consists of the accumulated data $\mathcal{D}_t$ and the competitive benchmark f_t^{EA}, which together define the GP posterior belief over $f_i(x)$. This posterior induces a distribution μ_t over the competition-aware payoff $Z_i^{\text{comp}}(x) = u(f_i(x), f_t^{\text{EA}})$. Robustness enters by allowing an adversary to distort μ_t to an alternative measure $\nu \ll \mu_t$, penalized by the KL divergence, yielding the robust evaluation $\mathcal{R}_\lambda$ and the resulting acquisition $\alpha_i^{\text{REA}}(x) = \mathcal{R}_\lambda(Z_i^{\text{comp}}(x))$. The next query is selected by maximizing $\alpha_i^{\text{REA}}(x)$ over $x \in \mathcal{X}$.

3.3. Algorithm Implementation

For the EA-style payoff $Z_i^{\text{comp}}(x) = (f_i(x) - f_t^{\text{EA}})^+$, we evaluate the expectation in Eq. (15) using Monte Carlo sampling from the one-dimensional GP posterior marginal. Although this expectation can be evaluated analytically for the rectified Gaussian payoff, the Monte Carlo estimator is retained because it applies directly to more general benchmark-dependent payoff transformations $u(\cdot,\cdot)$.

At iteration t and candidate $x \in \mathcal{X}$, the GP posterior marginal is

$$f_i(x) \mid \mathcal{D}_t \sim \mathcal{N}\big(\mu_t(x),\, \sigma_t^2(x)\big). \tag{19}$$

Draw S samples $Y^{(s)} \sim \mathcal{N}(\mu_t(x), \sigma_t^2(x))$ and map them to payoff samples $Z^{(s)} = \big(Y^{(s)} - f_t^{\text{EA}}\big)^+$. The expectation in (15) is approximated by Monte Carlo sampling. In the reported experiments, we use a fixed Monte Carlo size $S = 200$ for all RCBO acquisition evaluations. The same value of S is used across benchmark functions and random seeds to ensure comparable acquisition estimates.

$$\mathbb{E}_{\mu_t}\left[\exp\big(-Z_i^{\text{comp}}(x)/\lambda\big)\right] \approx \frac{1}{S}\sum_{s=1}^{S} \exp\big(-Z^{(s)}/\lambda\big), \tag{20}$$

yielding the Monte Carlo acquisition estimate

$$\widehat{\alpha}_i^{\text{REA}}(x;\lambda) = -\lambda \log\left(\frac{1}{S}\sum_{s=1}^{S} \exp\big(-Z^{(s)}/\lambda\big)\right). \tag{21}$$

The acquisition function is optimized over a finite candidate set rather than by gradient-based optimization. Specifically, at iteration t, we construct a candidate set $\mathcal{X}_{\text{cand},t} \subset \mathcal{X}$, evaluate $\widehat{\alpha}_i^{\text{REA}}(x;\lambda)$ for every $x \in \mathcal{X}_{\text{cand},t}$, and select the candidate with the largest acquisition value. Therefore, gradients of the Monte Carlo estimator are not used. This avoids instability that could otherwise arise from differentiating a stochastic acquisition estimate.

Algorithm 1: Robust Competitive Bayesian Optimization (agent i)

Input: Initial data $\mathcal{D}_{i,n_0}$; Monte Carlo size S; evaluation budget T; benchmark update rule for f_t^{EA}; adaptive risk-sensitivity rule for λ_t
. **Output:** Sequence of queried designs $\{x_{it}\}_{t=n_0+1}^{T}$.
Fit GP posterior on $\mathcal{D}_{i,n_0}$;
for $t = n_0, \ldots, T-1$ **do**
  Update GP posterior mean $\mu_{it}(\cdot)$ and variance $\sigma_{i,t}^2(\cdot)$ using $\mathcal{D}_{i,t}$ via (2)–(3);
  Update competitive benchmark f_{it}^{EA};
  Define the acquisition estimator $\widehat{\alpha}_i^{\text{REA}}(x;\lambda)$ via Monte Carlo approximation;
  Construct a finite candidate set $\mathcal{X}_{\text{cand},t} \subset \mathcal{X}$;
  Select next query:

$$x_{i,t+1} \in \arg\max_{x \in \mathcal{X}_{\text{cand},t}} \widehat{\alpha}_i^{\text{REA}}(x;\lambda);$$

  Evaluate $y_{i,t+1} = f_i(x_{i,t+1}) + \varepsilon_{i,t+1}$ with $\varepsilon_{i,t+1} \sim \mathcal{N}(0, v^2)$;
  Update dataset: $\mathcal{D}_{i,t+1} = \mathcal{D}_{i,t} \cup \{(x_{i,t+1}, y_{i,t+1})\}$;

Algorithm 1 summarizes the RCBO implementation for agent i. The benchmark f_t^{EA} is treated as an external competition state and updated from the current information available at iteration t (e.g., the opponent's best-so-far value). The next query is selected by maximizing the Monte Carlo estimate (21).

4. EXPERIMENTS

We evaluate RCBO against the nominal competitive Bayesian optimization baseline, i.e., CBO instantiated with the *Gap-Aware Threshold*(GAT) strategy. The GAT strategy is an adaptive competitive acquisition rule introduced in our prior CBO work [17], as summarized in Appendix A. In that framework, the improvement benchmark is dynamically adjusted according to the current performance gap between competing agents, so that the sampling strategy becomes more exploratory when an agent is behind and more exploitative when it is ahead. This mechanism enables the acquisition function to adapt its exploration–exploitation behavior to the evolving competitive state. When robustness is not considered (i.e., under risk-neutral evaluation), the proposed RCBO framework reduces to this nominal CBO-GAT strategy. Therefore, in the experiments we compare RCBO against CBO-GAT as the appropriate competition-aware baseline. We do not compare against classical single-agent BO, since the advantage of CBO over standard BO has already been empirically established in our prior work [17].

4.1. Experimental Settings

4.1.1. Objective functions and BO settings. We evaluate the proposed method on five benchmark functions: Cosines [25],

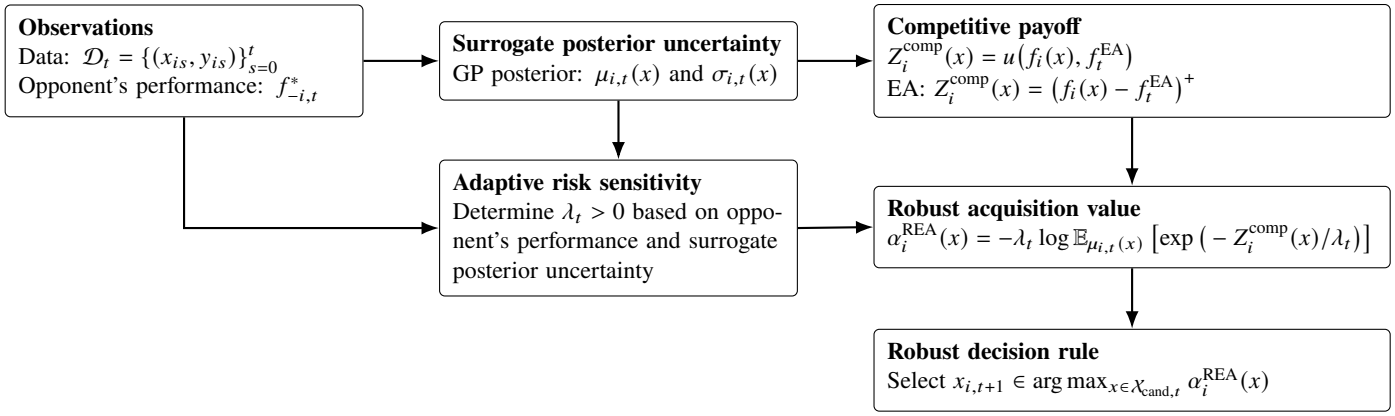


FIGURE 1: framework of robust competitive Bayesian optimization.

Alpine N.1 [26], Cross-in-tray [27], Schwefel [28], and Eggholder [27]. These functions are ordered approximately from simpler to more complex landscapes. In all experiments, each canonical domain is linearly rescaled to $[0, 1]^2$, and function values are standardized to have zero mean and unit variance for comparability. This rectangular-domain setting is used only for controlled benchmark evaluation. The acquisition rule itself only requires a feasible set of candidate designs over which the acquisition function can be evaluated. Detailed information for each benchmark function is provided in Table 1.

The benchmark functions are deterministic response surfaces. Stochasticity enters only through the additive Gaussian observation-noise model in Eq. (1) and through the GP posterior uncertainty induced by finite data. In this setting, the experimental uncertainty variance v^2 is treated as an independent nugget term added to the diagonal of the GP covariance matrix. Thus, the reported experiments assume additive homoscedastic observation noise, not multiplicative observation uncertainty, input-dependent noise, or correlated experimental errors. Therefore, the experiments evaluate robustness to noisy and uncertain acquisition evaluation rather than to an explicit time-varying structural shift of the objective function.

For all methods, the same GP modeling procedure is used. Input variables are linearly rescaled to $[0, 1]^2$, response values are standardized, and GP hyperparameters are re-estimated by log-marginal-likelihood maximization at each BO iteration [22]. Thus, differences between RCBO and the baseline are attributable to the robust competitive acquisition rule rather than to different surrogate-model training procedures.

For acquisition optimization, we use finite candidate-set maximization rather than gradient-based optimization. In each BO iteration, the candidate set $\mathcal{X}_{\mathrm{cand},t}$ is generated by Latin hypercube sampling in the normalized design domain, with $N_{\mathrm{cand}} = 5$ candidate points. RCBO and CBO-GAT evaluate their acquisition functions on the same candidate set within each random trial, and the point with the largest acquisition value is selected. Therefore, gradients of the Monte Carlo acquisition estimator are not used.

4.1.2. Adaptive risk sensitivity parameter. The expected-advantage payoff $Z_t(x)$ changes scale during optimization because its reference value (the incumbent benchmark) is updated online. Consequently, using a fixed absolute risk parameter λ is not comparable across iterations or across objective functions: the entropic robust evaluation in Eq. (15) depends on λ only through the dimensionless ratio Z/λ. To maintain comparable robustness while tracking the evolving payoff magnitude, we introduce the scale-tracking parameterization

$$\lambda_t = \kappa_t s_t, \qquad s_t := \mathrm{IQR}\big(Z_t(\mathcal{X})\big), \tag{22}$$

where $\mathcal{X}$ is the discrete candidate set used for acquisition maximization (i.e., the finite set of candidate design points over which the acquisition function is maximized) and $Z_t(\mathcal{X}) = \{Z_t(x) : x \in \mathcal{X}\}$. The interquartile range provides a robust estimate of the payoff dispersion, ensuring that the risk parameter adapts to the scale of the competitive payoff.

The key mechanism is illustrated in Fig. 2. The effective risk sensitivity is determined by two interacting quantities: surrogate posterior uncertainty, summarized by the GP posterior standard deviation over the candidate set, and the competitive standing relative to the opponent. Accordingly, we decompose the temperature parameter as

$$\kappa_t = \kappa_t^{\mathrm{model}} \kappa_t^{\mathrm{opp}}, \tag{23}$$

where each component controls robustness from a different perspective.

1) Surrogate posterior uncertainty. The left branch of Fig. 2 describes how surrogate posterior uncertainty influences risk sensitivity. When the GP posterior variance is large over the candidate set, exploration decisions based solely on the posterior mean may be unreliable. In this regime the algorithm adopts a more cautious strategy by reducing the temperature and therefore increasing robustness (i.e., producing a more conservative, risk-averse evaluation of the competitive payoff). As the model becomes more confident, robustness is gradually relaxed.

To quantify this effect, we measure the surrogate posterior uncertainty over the candidate set by

$$u_t := \mathrm{median}_{x \in \mathcal{X}} \sigma_t(x). \tag{24}$$

Here, $\sigma_t(x)$ denotes the posterior standard deviation of the latent response $f_i(x)$, not the additive observation-noise standard deviation v. Thus, u_t is used only as a scalar summary of the

TABLE 1: Definitions of the five benchmarks with known global minima and canonical domains. In all experiments, domains are linearly rescaled to $[0, 1]^2$ and function values standardized.

Name	Formula	Global minimum	Canonical domain
Cosines	$f(\mathbf{x}) = 1 - \frac{1}{2}\sum_{i=1}^{2} x_i^2 - 0.3\sum_{i=1}^{2}\cos(3\pi x_i)$	$f(0.314,\ 0.303) \approx -1.596$	$x_i \in [-1,\ 1]$
Alpine N.1	$f(x, y) = \lvert x \sin x + 0.1x\rvert + \lvert y \sin y + 0.1y\rvert$	$f(0,\ 0) = 0$	$x, y \in [0,\ 10]$
Cross-in-tray	$f(x, y) = -10^{-4}\left(\lvert\sin x\ \sin y\ \exp\left(\lvert 100 - \frac{\sqrt{x^2+y^2}}{\pi}\rvert\right)\rvert + 1\right)^{0.1}$	$f(\pm 1.34941,\ \pm 1.34941) \approx -2.06261$	$x, y \in [-10,\ 10]$
Schwefel	$f(x, y) = 2 \cdot 418.9829 - \left[x \sin\sqrt{\lvert x\rvert} + y \sin\sqrt{\lvert y\rvert}\right]$	$f(420.9687,\ 420.9687) = 0$	$x, y \in [-500,\ 500]$
Eggholder	$f(x, y) = -(y + 47)\ \sin\left(\sqrt{\lvert x/2 + y + 47\rvert}\right)\ -\ x\ \sin\left(\sqrt{\lvert x - (y + 47)\rvert}\right)$	$f(512,\ 404.2319) \approx -959.6407$	$x, y \in [-512,\ 512]$

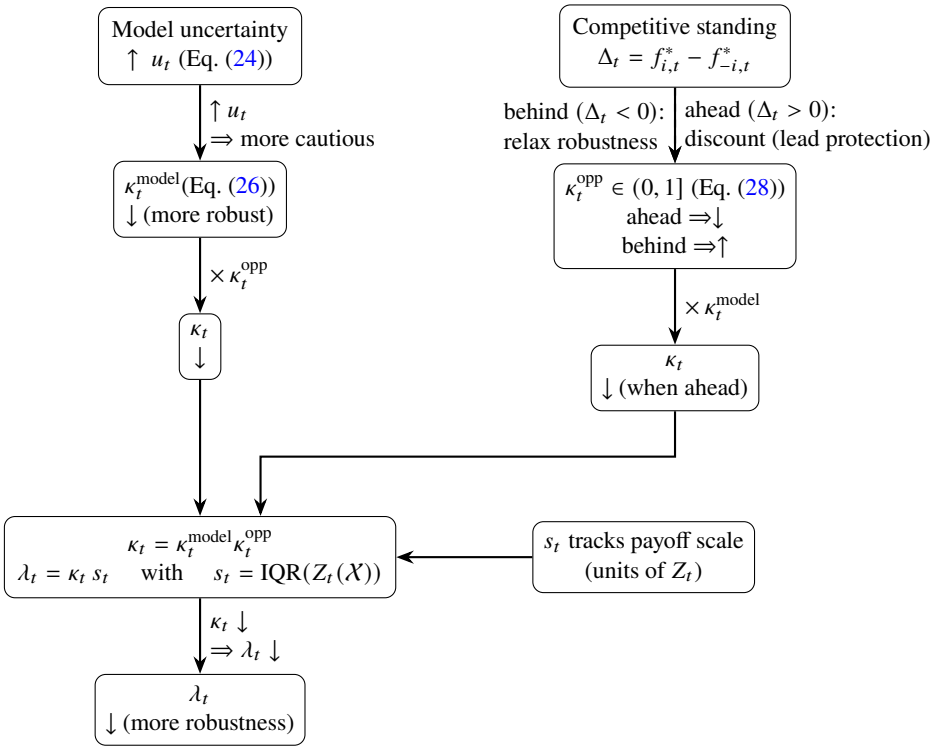


FIGURE 2: Logic flow of the adaptive risk sensitivity: model uncertainty and competitive standing jointly determine $\kappa_t = \kappa_t^{\text{model}} \kappa_t^{\text{opp}}$, which together with payoff-scale tracking s_t sets $\lambda_t = \kappa_t s_t$.

surrogate model's current uncertainty over candidate designs. A large value of u_t indicates that the model is still uncertain across the candidate set, corresponding to the "more cautious" regime in Fig. 2. We normalize this quantity into a progress variable

$$g_t := \text{clip}\left(\frac{u_0 - u_t}{u_0 - u_{\min}},\ 0,\ 1\right), \qquad u_{\min} := \eta u_0, \tag{25}$$

where $\text{clip}(z, 0, 1)$ truncates the value z to the interval $[0, 1]$, i.e., values below 0 are set to 0 and values above 1 are set to 1. The quantity u_0 denotes the initial surrogate uncertainty level at the beginning of optimization (computed from Eq. (24)), while $\eta \in (0, 1)$ is a small constant that defines the minimum uncertainty level $u_{\min}$ used for normalization. As the optimization progresses and the surrogate becomes more confident (u_t decreases), the normalized progress variable g_t increases from 0 toward 1. The model-driven temperature is then defined by

$$\kappa_t^{\text{model}} = \frac{1}{\max\{1 - g_t, \varepsilon\}}, \tag{26}$$

which produces the trend depicted in the left flow of Fig. 2: higher uncertainty ($u_t \uparrow$) leads to a smaller temperature ($\kappa_t \downarrow$) and therefore a smaller risk parameter ($\lambda_t \downarrow$), corresponding to stronger robustness (i.e., a more risk-averse evaluation of the competitive payoff) As uncertainty decreases during optimization, κ_t^{model} increases and the algorithm gradually approaches risk-neutral behavior.

2) Competitive standing. The right branch of Fig. 2 describes how the current competitive position affects risk sensitivity. Let

$$\Delta_t := f^*_{i,t} - f^*_{-i,t}, \qquad \tilde{\Delta}_t := \Delta_t / s_{\Delta,t}, \tag{27}$$

where $s_{\Delta,t}$ is a robust scale used to normalize the performance gap. The opponent-driven component is defined as

$$\kappa_t^{\text{opp}} = \text{sigmoid}(-\tilde{\Delta}_t) = \frac{1}{1 + \exp(\tilde{\Delta}_t)}, \tag{28}$$

where $\text{sigmoid}(z) = 1/(1 + e^{-z})$ denotes the logistic sigmoid function. which lies in $(0, 1)$. This formulation produces the behavior illustrated in the right flow of Fig. 2. When the agent is ahead ($\Delta_t > 0$), the temperature is discounted ($\kappa_t^{\text{opp}} \downarrow$), reducing

TABLE 2: Convergence speed across 50 random seeds. Values report the mean number of evaluations required for Agent A to reach a predefined performance threshold, with standard deviations shown in parentheses; lower values indicate faster convergence. Agent A uses RCBO and Agent B uses the CBO-GAT baseline.

Function	CBO-GAT	RCBO	Improvement
Cosines	16.16(5.36)	**15.56(4.77)**	3.7%
Alpine N.1	29.76(5.37)	**27.92(4.67)**	6.2%
Cross-in-tray	161.80(42.50)	**141.80(32.30)**	12.4%
Schwefel	256.90(89.10)	**233.36(54.38)**	9.2%
Eggholder	552.00(63.29)	**480.60(50.20)**	12.9%

κ_t and therefore λ_t; this increases robustness and protects the current lead. Conversely, when the agent is behind ($\Delta_t < 0$), the discount is relaxed ($\kappa_t^{\text{opp}} \uparrow$), allowing a larger temperature and encouraging more aggressive exploration to close the gap.

Combining these two components yields the overall temperature κ_t, which together with the payoff-scale factor s_t determines the adaptive risk parameter λ_t. As illustrated in Fig. 2, the model uncertainty controls the long-term annealing of robustness, while the opponent-driven factor adjusts the temperature dynamically according to the current competitive advantage.

4.2. Experimental Results

4.2.1. Convergence speed. Let us define the scenarios where Agent A is competing against Agent B in finding the global optimum of the five benchmark functions following the RCBO framework. Table 2 reports the average number of evaluations required for Agent A to reach a predefined performance threshold across 50 random seeds (lower values indicate faster convergence). Across all benchmark functions, RCBO consistently requires fewer evaluations than the CBO-GAT baseline, demonstrating that the proposed robustness mechanism improves convergence speed.

On relatively smooth landscapes such as Cosines and Alpine N.1, both methods converge rapidly, but RCBO still reaches the target performance slightly earlier on average. The standard deviations are also comparable or smaller for RCBO, indicating stable behavior across random seeds. The advantage becomes more pronounced on more challenging multimodal functions—including Cross-in-tray, Schwefel, and Eggholder—where RCBO achieves noticeably faster convergence. Table 2 also reports the percentage improvement of RCBO over the CBO-GAT baseline, highlighting the relative reduction in the number of evaluations required to reach the target performance. For example, on Cross-in-tray the average evaluation count decreases from 161.80 to 141.80, and on Eggholder from 552.00 to 480.60. These improvements suggest that the adaptive risk mechanism helps guide the search more effectively under high uncertainty and rugged landscapes. Overall, the results indicate that incorporating distributional robustness does not slow down optimization. Instead, it can accelerate convergence by stabilizing decision-making in the early stages of exploration and reducing inefficient evaluations.

4.2.2. Strategic performance. Beyond convergence speed, we analyze the head-to-head competitive dynamics between the two agents throughout the optimization process. In this setup, Agent A implements the proposed RCBO framework, while Agent B follows the nominal CBO-GAT strategy. Specifically, we track which agent leads at each iteration and by how much across random seeds. At iteration t, let $f_A^{*,s}(t)$ and $f_B^{*,s}(t)$ denote the incumbent best values of Agent A and Agent B for seed s. We compute the seed-averaged win probability (Win Rate) and the average performance advantage:

$$\widehat{\mathrm{WR}}_A(t) = \frac{1}{S}\sum_{s=1}^{S} W_A^s(t), \tag{29}$$

$$\overline{\Delta}_A(t) = \frac{1}{S}\sum_{s=1}^{S}\left(f_A^{*,s}(t) - f_B^{*,s}(t)\right), \tag{30}$$

where the instantaneous outcome indicator is defined as

$$W_A^s(t) = \begin{cases} 1, & f_A^{*,s}(t) > f_B^{*,s}(t), \\ \frac{1}{2}, & f_A^{*,s}(t) = f_B^{*,s}(t), \\ 0, & f_A^{*,s}(t) < f_B^{*,s}(t). \end{cases} \tag{31}$$

Thus $\widehat{\mathrm{WR}}_A(t) \in [0,1]$ measures the probability that Agent A is ahead at iteration t, while $\overline{\Delta}_A(t)$ quantifies the average performance gap. To summarize leadership over the entire evaluation horizon $t = 1, \ldots, T$, we compute two time-averaged leadership fractions. The first measures the fraction of the budget during which Agent A leads strictly, while the second measures periods of parity:

$$\mathrm{TLF}_{\mathrm{win}} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{\widehat{\mathrm{WR}}_A(t) > 0.5\}, \tag{32}$$

$$\mathrm{TLF}_{\mathrm{tie}} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{\widehat{\mathrm{WR}}_A(t) = 0.5\}. \tag{33}$$

Both metrics lie in $[0,1]$. A larger $\mathrm{TLF}_{\mathrm{win}}$ indicates that Agent A leads for a larger portion of the optimization horizon, while a larger $\mathrm{TLF}_{\mathrm{tie}}$ indicates more frequent parity between the competing agents.

Table 3 summarizes the time-leadership fractions for the benchmark problems. Across all functions, RCBO maintains leadership for a larger portion of the evaluation horizon relative to the CBO-GAT baseline. On smoother landscapes such as Cosines and Alpine N.1, the advantage is moderate: RCBO leads for approximately 52% and 55% of the optimization horizon, respectively. These problems also exhibit relatively large tie fractions (32% and 28%), indicating that both algorithms frequently reach similar incumbent solutions once convergence begins. The advantage becomes more pronounced on more challenging multimodal functions. On Cross-in-tray and Schwefel, RCBO leads for 63% and 60% of the evaluation horizon, respectively, while the tie fractions decrease to roughly 22–24%. This suggests that RCBO maintains a clearer competitive edge during the search process. The strongest difference appears on the highly rugged

TABLE 3: Fraction of the evaluation budget during which Agent A leads (TLF_{win}) or ties (TLF_{tie}). Agent A uses RCBO and Agent B uses the CBO-GAT baseline.

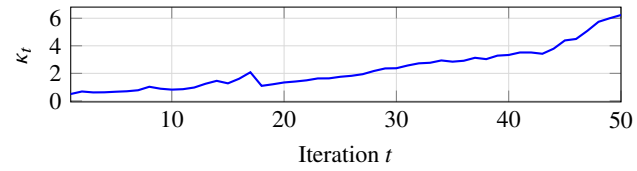
Function	TLF_{win}	TLF_{tie}
Cosines	0.52	0.32
Alpine N.1	0.55	0.28
Cross-in-tray	0.63	0.22
Schwefel	0.60	0.24
Eggholder	0.66	0.21

Eggholder function, where RCBO leads for 66% of the optimization horizon with a tie fraction of only 21%. This result indicates that the robust competitive acquisition allows the agent to sustain a strategic advantage more consistently in difficult optimization landscapes.

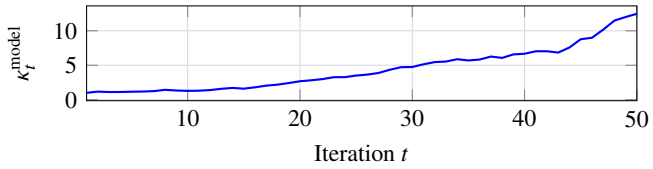
4.2.3. Adaptive risk dynamics. To better understand how the adaptive risk mechanism behaves during optimization, we visualize the evolution of the temperature parameter κ_t and its two components on representative benchmark functions. Figures 3 and 4 show these trajectories for the Cosines function, which has a relatively smooth landscape, and the Eggholder function, which is highly multimodal. The combined temperature κ_t (subfigures (a)) determines the overall level of robustness, while κ_t^{model} (subfigures (b)) captures the influence of model uncertainty, and κ_t^{opp} (subfigures (c)) reflects adjustments based on the current competitive standing.

Two observations emerge from these figures. *First*, the overall temperature κ_t is primarily driven by the model-driven component. As shown in Figs. 3(b) and 4(b), κ_t^{model} increases over iterations as posterior uncertainty decreases, which causes the combined temperature in Figs. 3(a) and 4(a) to gradually rise and anneal toward a more risk-neutral regime. Meanwhile, the opponent-driven component κ_t^{opp} fluctuates during the early and middle stages of optimization as the competitive gap changes, but eventually stabilizes around 0.5 in Figs. 3(c) and 4(c) once both agents approach similar incumbent values and the performance gap becomes negligible.

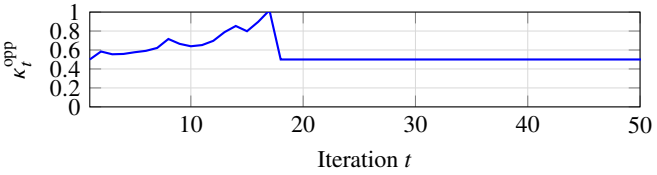
Second, the rate at which κ_t^{model} grows differs substantially between simple and complex landscapes. For the smoother Cosines function (Fig. 3), κ_t^{model} increases rapidly, indicating that the surrogate model quickly reduces uncertainty and the algorithm transitions to near risk-neutral behavior within relatively few iterations. In contrast, for the highly multimodal Eggholder function (Fig. 4), κ_t^{model} grows much more slowly and remains comparatively smaller even after many evaluations. This indicates that the surrogate model retains substantial epistemic uncertainty for a longer period on the rugged landscape. Consequently, the adaptive risk mechanism keeps the temperature relatively low and maintains stronger robustness during the search. This behavior highlights the advantage of RCBO: by adjusting risk sensitivity according to model uncertainty, the algorithm preserves a cautious evaluation strategy in difficult optimization problems where premature risk-neutral decisions could lead to unstable or misleading improvements.



(a) Combined temperature κ_t.

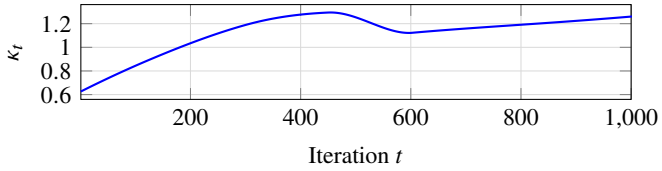


(b) Model-driven component κ_t^{model}.

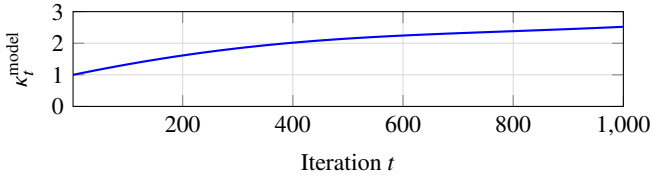


(c) Opponent-driven discount κ_t^{opp}.

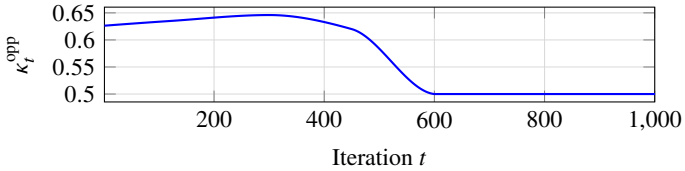
FIGURE 3: Cosines function: evolution of the adaptive temperature κ_t and its components.



(a) Combined temperature κ_t.



(b) Model-driven temperature component κ_t^{model}.



(c) Opponent-driven discount κ_t^{opp}.

FIGURE 4: Eggholder function: evolution of the adaptive temperature κ_t and its components. Early iterations exhibit conservative behavior driven by model uncertainty, while later iterations approach risk-neutral behavior as uncertainty decreases. The opponent component discounts the temperature when ahead and relaxes robustness when behind.

5. DISCUSSION

This section synthesizes the main findings of the experimental study and interprets them in the context of robust competitive decision-making. Rather than focusing only on performance

differences between RCBO and the CBO-GAT baseline, the discussion highlights what the results reveal about how uncertainty and competitive standing jointly shape risk-sensitive acquisition decisions in competitive Bayesian optimization (BO). The experimental results provide several insights into how the proposed adaptive risk mechanism operates in competitive BO.

First, problem complexity strongly influences the evolution of risk sensitivity. Figures 3 and 4 show that the combined temperature κ_t evolves smoothly for the Cosines function but exhibits larger fluctuations for the highly multimodal Eggholder function. On the simpler landscape, κ_t increases gradually as the surrogate model becomes more confident. In contrast, the Eggholder function displays sharper transitions in κ_t, particularly during the early and middle stages of optimization. This behavior reflects the stronger epistemic uncertainty and irregular improvement patterns associated with rugged landscapes. From an engineering decision-making perspective, this result highlights the importance of adapting robustness to landscape complexity: smoother problems allow faster transition toward risk-neutral decisions, whereas complex design spaces benefit from sustained cautious exploration with adaptive risk attitude.

Second, model uncertainty primarily governs the long-term evolution of robustness. As shown in Figs. 3(b) and 4(b), the model-driven component κ_t^{model} increases monotonically over iterations as posterior uncertainty decreases. This trend produces the annealing behavior observed in the combined temperature κ_t in Fig. 3(a) and Fig. 4(a). However, the rate of increase differs substantially between functions. On the Cosines benchmark, uncertainty collapses rapidly and κ_t^{model} rises quickly, allowing the algorithm to approach a risk-neutral regime early. In contrast, for the Eggholder function κ_t^{model} grows much more slowly, indicating persistent uncertainty in the rugged search landscape. This mechanism ensures that robustness is maintained longer in difficult optimization environments, allowing the algorithm to avoid overly confident decisions when the surrogate model is still uncertain.

Third, competitive standing introduces strategic adjustments to risk sensitivity. The opponent-driven component κ_t^{opp} responds dynamically to the incumbent performance gap between agents. As illustrated in Figs. 3(c) and 4(c), the value of κ_t^{opp} decreases when the agent is ahead and increases when the agent falls behind. This behavior discounts the temperature when a lead is established, protecting the incumbent advantage through more conservative decisions. Conversely, when the agent trails the opponent, robustness is relaxed to encourage more aggressive exploration. In engineering design scenarios where solutions are evaluated relative to competitors or benchmarks, such strategic adaptation helps balance the exploitation of current advantages with the exploration needed to recover from temporary disadvantages.

Fourth, the joint adaptation of uncertainty and competitive state improves decision stability. The results across convergence speed, strategic performance, and adaptive risk dynamics indicate that combining model uncertainty and opponent performance provides a principled mechanism for regulating robustness during optimization. When uncertainty is high or the agent is ahead, the algorithm adopts a more cautious strategy. When confidence increases or the agent falls behind, it gradually becomes more risk-neutral. This joint adaptation stabilizes early-stage exploration while maintaining responsiveness to competitive pressure. As demonstrated in the experiments, this mechanism allows RCBO to achieve faster convergence and maintain competitive leadership across multiple benchmark functions. More broadly, these findings suggest that incorporating both epistemic uncertainty and strategic context into acquisition design can improve reliability in engineering optimization problems where decisions must be made under both uncertainty and competition.

6. CONCLUSION

This paper introduced Robust Competitive Bayesian Optimization (RCBO), a principled extension of competitive Bayesian optimization that integrates distributional robustness into competition-aware acquisition decisions. By evaluating competitive advantage under a KL-distorted posterior, RCBO incorporates ambiguity aversion into the acquisition process while preserving the tractability of Gaussian-process–based Bayesian optimization. To maintain consistent risk sensitivity across iterations, we further proposed an adaptive scaling mechanism that adjusts robustness according to both model uncertainty and competitive standing. Empirical results on a range of benchmark functions demonstrate that RCBO improves convergence efficiency, maintains competitive leadership for a larger portion of the optimization horizon, and stabilizes early-stage decision-making under high uncertainty. The results provide several practical insights for engineering optimization under competition. First, model uncertainty plays a dominant role in determining long-term risk sensitivity, naturally inducing an annealing behavior as the surrogate model becomes more confident. Second, competitive standing introduces strategic adjustments that balance lead protection with recovery from disadvantage. Together, these mechanisms allow RCBO to regulate robustness dynamically, improving reliability in sequential design problems where decisions must be made under both uncertainty and competitive pressure.

Several limitations suggest directions for future research. First, the present study focuses on two-agent competition with a scalar benchmark representing the opponent's best-so-far performance. Many real engineering environments involve multiple interacting agents or more complex competitive structures. Extending RCBO to multi-agent settings or to richer game-theoretic interaction models would allow the framework to capture more realistic strategic dynamics. Second, this work primarily provides empirical validation of the proposed robust competitive acquisition mechanism. Developing theoretical guarantees—such as regret bounds, equilibrium properties, or convergence analysis under robust competitive dynamics—remains an important direction for future work and would further strengthen the connection between robust decision theory (e.g., [29]) and competitive Bayesian optimization.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support from the National Science Foundation through the grants CMMI-2321463.

REFERENCES

[1] Li, Changhe, Han, Shoufei, Zeng, Sanyou and Yang, Shengxiang. "Robust optimization." *Intelligent Optimization: Principles, Algorithms and Applications*. Springer (2024): pp. 239–251.

[2] Rahimian, Hamed and Mehrotra, Sanjay. "Distributionally robust optimization: A review." *arXiv preprint arXiv:1908.05659* (2019).

[3] Wald, Abraham. *Statistical Decision Functions*. Wiley (1950).

[4] Gilboa, Itzhak and Schmeidler, David. "Maxmin Expected Utility with Non-Unique Prior." *Journal of Mathematical Economics* Vol. 18 No. 2 (1989): pp. 141–153. DOI 10.1016/0304-4068(89)90018-9.

[5] Hansen, Lars Peter and Sargent, Thomas J. *Robustness*. Princeton University Press (2008).

[6] Maccheroni, Fabio, Marinacci, Massimo and Rustichini, Aldo. "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences." *Econometrica* Vol. 74 No. 6 (2006): pp. 1447–1498. DOI 10.1111/j.1468-0262.2006.00705.x.

[7] Ben-Tal, Aharon, den Hertog, Dick, De Waegenaere, Anja, Melenberg, Bertrand and Rennen, Gijs. "Robust Solutions of Optimization Problems Affected by Uncertain Probabilities." *Management Science* Vol. 59 No. 2 (2013): pp. 341–357. DOI 10.1287/mnsc.1120.1641.

[8] Rahimian, Hamed and Mehrotra, Sanjay. "Distributionally Robust Optimization: A Review." *arXiv preprint arXiv:1908.05659* (2019): pp. 1–64URL 1908.05659.

[9] Mockus, Jonas, Tiesis, Vytautas and Zilinskas, Antanas. "The Application of Bayesian Methods for Seeking the Extremum." Dixon, L. C. W. and Szegö, G. P. (eds.). *Towards Global Optimization*: pp. 117–129. 1975. Elsevier.

[10] Jones, Donald R., Schonlau, Matthias and Welch, William J. "Efficient Global Optimization of Expensive Black-Box Functions." *Journal of Global Optimization* Vol. 13 No. 4 (1998): pp. 455–492.

[11] Shahriari, Bobak, Swersky, Kevin, Wang, Ziyu, Adams, Ryan P. and de Freitas, Nando. "Taking the Human Out of the Loop: A Review of Bayesian Optimization." *Proceedings of the IEEE* Vol. 104 No. 1 (2016): pp. 148–175.

[12] Frazier, Peter I. "A Tutorial on Bayesian Optimization." *arXiv preprint arXiv:1807.02811* (2018): pp. 1–22URL 1807.02811.

[13] Yue, Yisong, Broder, Josef, Kleinberg, Robert and Joachims, Thorsten. "The K-armed Dueling Bandits Problem." *Proceedings of the 22nd Annual Conference on Learning Theory (COLT)*: pp. 1–28. 2009. JMLR Workshop and Conference Proceedings.

[14] Zoghi, Masrour, Whiteson, Shimon, Munos, Rémi and de Rijke, Maarten. "Relative Upper Confidence Bound for the K-armed Dueling Bandit Problem." *Proceedings of the 31st International Conference on Machine Learning (ICML)*: pp. 10–18. 2014. PMLR.

[15] Fudenberg, Drew and Tirole, Jean. *Game Theory*. MIT Press, Cambridge, MA (1991).

[16] Littman, Michael L. "Markov Games as a Framework for Multi-Agent Reinforcement Learning." *Proceedings of the 11th International Conference on Machine Learning (ICML)*: pp. 157–163. 1994. Morgan Kaufmann.

[17] Chen, Siyu, Bayrak, Alparslan Emrah and Sha, Zhenghui. "Competitive Bayesian Optimization." *PNAS Nexus* (2025). Under review.

[18] Srinivas, Niranjan, Krause, Andreas, Kakade, Sham M. and Seeger, Matthias. "Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design." *Proceedings of the 27th International Conference on Machine Learning (ICML)*: pp. 1015–1022. 2010. Omnipress.

[19] Bogunovic, Ilija, Scarlett, Jonathan, Krause, Andreas and Cevher, Volkan. "Adversarially Robust Optimization with Gaussian Processes." *Advances in Neural Information Processing Systems (NeurIPS)*, Vol. 31: pp. 5760–5770. 2018. Curran Associates, Inc.

[20] Kirschner, Johannes, Mutny, Mojmir, Hiller, Nico, Ischebeck, Rainer and Krause, Andreas. "Distributionally Robust Bayesian Optimization." *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS)*, Vol. 108: pp. 2174–2184. 2020. PMLR.

[21] Donsker, Monroe D. and Varadhan, S. R. Srinivasa. "Asymptotic Evaluation of Certain Markov Process Expectations for Large Time." *Communications on Pure and Applied Mathematics* Vol. 28 No. 1 (1975): pp. 1–47.

[22] Rasmussen, Carl Edward and Williams, Christopher K. I. *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA (2006).

[23] Dembo, Amir and Zeitouni, Ofer. *Large Deviations Techniques and Applications*, 2nd ed. Springer (1998).

[24] Föllmer, Hans and Schied, Alexander. "Convex measures of risk and trading constraints." *Finance and Stochastics* Vol. 6 No. 4 (2002): pp. 429–447.

[25] Ali, M. Montaz, Khompatraporn, Charoenchai and Zabinsky, Zelda B. "A Numerical Evaluation of Several Stochastic Algorithms on Selected Continuous Global Optimization Test Problems." *Journal of Global Optimization* Vol. 31 No. 4 (2005): pp. 635–672. DOI 10.1007/s10898-004-9972-2.

[26] Molga, Marcin and Smutnicki, Czesław. "Test Functions for Optimization Needs." Technical report no. Institute of Computer Engineering, Control and Robotics, Wrocław University of Technology. 2005. URL https://robertmarks.org/Classes/ENGR5358/Papers/functions.pdf. Technical report.

[27] Surjanovic, Sonja and Bingham, Derek. "Cross-in-Tray Function." Virtual Library of Simulation Experiments: Test Functions and Datasets, Simon Fraser University (2013). Accessed 2025-09-30.

[28] Schwefel, Hans-Paul. *Numerical Optimization of Computer Models*. John Wiley & Sons, Chichester, UK (1981).

[29] Maccheroni, Fabio, Marinacci, Massimo and Rustichini, Aldo. "Ambiguity aversion, robustness, and the variational representation of preferences." *Econometrica* Vol. 74 No. 6 (2006): pp. 1447–1498.

[30] Komorowski, Michal, Marshall, Darren C., Salciccioli, Joshua D. and Crutain, Yankel. "Exploratory Data Analysis." *Secondary Analysis of Electronic Health Records*. Springer, Cham (2016): Chap. 15, pp. 185–203. DOI 10.1007/978-3-319-43742-2_15.

[31] Cover, Thomas M. and Thomas, Joy A. *Elements of Information Theory*, 2nd ed. Wiley (2006).

[32] Brochu, Eric, Cora, Vlad M. and de Freitas, Nando. "A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning." Technical report. University of British Columbia. 2010.

[33] Shapley, Lloyd S. "Stochastic Games." *Proceedings of the National Academy of Sciences* Vol. 39 No. 10 (1953): pp. 1095–1100.

[34] Harsanyi, John C. "Games with Incomplete Information Played by Bayesian Players." *Management Science* Vol. 14 No. 3 (1967): pp. 159–182.

[35] Silver, David, Schrittwieser, Julian, Simonyan, Karen, Antonoglou, Ioannis, Huang, Aja, Guez, Arthur, Hubert, Thomas, Baker, Lucas, Lai, Matthew, Bolton, Adrian, Chen, Yutian, Lillicrap, Timothy, Hui, Fan, Sifre, Laurent, van den Driessche, George, Graepel, Thore and Hassabis, Demis. "Mastering the Game of Go without Human Knowledge." *Nature* Vol. 550 (2017): pp. 354–359.

[36] Bernstein, Daniel S., Givan, Robert, Immerman, Neil and Zilberstein, Shlomo. "The Complexity of Decentralized Control of Markov Decision Processes." *Mathematics of Operations Research* Vol. 27 No. 4 (2002): pp. 819–840.

[37] Oliehoek, Frans A. and Amato, Christopher. *A Concise Introduction to Decentralized POMDPs*. Springer (2016).

[38] Tan, Matthias H. Y. and Wu, C. F. J. "Robust Design Optimization with Quadratic Loss Derived from Gaussian Process Models." *Technometrics* Vol. 54 No. 1 (2012): pp. 51–63. DOI 10.1080/00401706.2012.648866.

[39] Feng, Zebiao, Wang, Jianjun, Ma, Yizhong and Tu, Yiliu. "Robust Parameter Design Based on Gaussian Process with Model Uncertainty." *International Journal of Production Research* Vol. 59 No. 9 (2021): pp. 2772–2788. DOI 10.1080/00207543.2020.1740344.

[40] Kennedy, Marc C. and O'Hagan, Anthony. "Bayesian Calibration of Computer Models." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* Vol. 63 No. 3 (2001): pp. 425–464. DOI 10.1111/1467-9868.00294.

APPENDIX A. OVERVIEW OF COMPETITIVE BAYESIAN OPTIMIZATION

Competitive Bayesian Optimization (CBO) extends classical Bayesian optimization (BO) to settings where decision quality is evaluated relative to an external benchmark or opponent. In many engineering and scientific applications, design teams or autonomous agents operate under limited evaluation budgets while competing to outperform a rival solution rather than simply improving an absolute objective value. Classical BO methods such as Expected Improvement (EI) measure progress relative to an agent's own best-observed value and therefore assume a self-referential notion of improvement. This formulation does not account for situations in which the success of a design depends on surpassing an opponent's current performance.

CBO addresses this limitation by introducing a competition-aware acquisition rule that evaluates candidate designs relative to a competitive benchmark. Consider two agents exploring the same design space. Each agent maintains a Gaussian-process surrogate model of the unknown objective function and observes the opponent's best-so-far performance. Instead of optimizing improvement relative to its own incumbent, the agent evaluates candidate points using an *Expected Advantage* (EA) acquisition function. EA generalizes Expected Improvement by replacing the self-referential improvement threshold with a benchmark that depends on both agents' current best-observed values.

Formally, let $f_i(x)$ denote the objective function for agent i, and let $f^*_{i,t}$ and $f^*_{-i,t}$ denote the best observed values for the agent and its opponent at iteration t. CBO defines a competitive benchmark

$$f_t^{\mathrm{EA}} = w\Big(f^*_{i,t},\, f^*_{-i,t}\Big),$$

where $w(\cdot)$ determines how the benchmark depends on the two agents' performances. The Expected Advantage acquisition function evaluates the expected improvement of a candidate design relative to this benchmark. This mechanism directs sampling toward designs that are most likely to improve the agent's competitive standing.

To adapt the exploration–exploitation trade-off to the competitive state, CBO further introduces a gap-aware thresholding mechanism that adjusts the benchmark based on the current performance gap between agents. When an agent is ahead, the threshold encourages reliable refinements that preserve the lead; when the agent is behind, the threshold promotes exploration of higher-upside regions that may overturn the opponent's advantage. This adaptive mechanism links the sampling policy directly to the evolving competitive state.

Through this competition-aware acquisition strategy, CBO provides a principled framework for sequential decision-making under competition. It preserves the sample efficiency of Bayesian optimization while enabling agents to dynamically adjust their search behavior in response to an opponent's progress.

APPENDIX B. VARIATIONAL REPRESENTATION WITH KL PENALTY

Theorem 1 (Variational representation)**.** *For any bounded random variable Z and $\lambda > 0$,*

$$\inf_{\nu\in\mathcal{P}} \Big\{\mathbb{E}_\nu[Z] + \lambda D_{\mathrm{KL}}(\nu\|\mu)\Big\} = -\lambda \log \mathbb{E}_\mu\big[e^{-Z/\lambda}\big]. \tag{34}$$

Moreover, the infimum is attained at the distorted measure $\nu^\star \in \mathcal{P}$ *defined by*

$$\nu_k^\star = \frac{\mu_k \exp(-Z_k/\lambda)}{\sum_j \mu_j \exp(-Z_j/\lambda)}, \qquad Z_k := Z(\omega_k). \tag{35}$$

Proof. Write the objective in (11) explicitly as

$$J(\nu) := \sum_k \nu_k Z_k + \lambda \sum_k \nu_k \log \frac{\nu_k}{\mu_k},$$

subject to $\sum_k \nu_k = 1$ and $\nu_k \geq 0$ for all k. This is a convex optimization problem over the probability simplex.

Introduce a Lagrange multiplier $\eta \in \mathbb{R}$ for the constraint $\sum_k \nu_k = 1$ and consider the Lagrangian

$$\mathcal{L}(\nu, \eta) = \sum_k \nu_k Z_k + \lambda \sum_k \nu_k \log \frac{\nu_k}{\mu_k} + \eta\Big(\sum_k \nu_k - 1\Big).$$

For interior optima with $\nu_k > 0$, the first-order condition $\partial \mathcal{L}/\partial \nu_k = 0$ gives

$$0 = \frac{\partial \mathcal{L}}{\partial \nu_k} = Z_k + \lambda\Big(\log \frac{\nu_k}{\mu_k} + 1\Big) + \eta$$
$$\implies \quad \log \frac{\nu_k}{\mu_k} = -\frac{Z_k}{\lambda} - 1 - \frac{\eta}{\lambda}.$$

Exponentiating both sides yields

$$\frac{\nu_k}{\mu_k} = \exp\Big(-\frac{Z_k}{\lambda} - 1 - \frac{\eta}{\lambda}\Big) = C \exp\Big(-\frac{Z_k}{\lambda}\Big), \tag{36}$$

where $C := \exp(-1 - \eta/\lambda)$ is a normalizing constant that does not depend on k.

Therefore

$$\nu_k = C\, \mu_k \exp\Big(-\frac{Z_k}{\lambda}\Big).$$

Imposing $\sum_k \nu_k = 1$ determines C:

$$1 = \sum_k \nu_k = C \sum_k \mu_k \exp\Big(-\frac{Z_k}{\lambda}\Big) \quad \implies \quad C = \frac{1}{\sum_j \mu_j \exp(-Z_j/\lambda)}.$$

Substituting back gives (35), so $\nu^\star$ is the unique optimizer.

To evaluate the minimum value $J(\nu^\star)$, we first compute the KL divergence at $\nu^\star$. From (36),

$$\log \frac{\nu_k^\star}{\mu_k} = -\frac{Z_k}{\lambda} - 1 - \frac{\eta}{\lambda}.$$

Hence

$$\begin{aligned} D_{\mathrm{KL}}(\nu^\star \| \mu) &= \sum_k \nu_k^\star \log \frac{\nu_k^\star}{\mu_k} \\ &= \sum_k \nu_k^\star \Big(-\frac{Z_k}{\lambda} - 1 - \frac{\eta}{\lambda}\Big) \\ &= -\frac{1}{\lambda} \sum_k \nu_k^\star Z_k - \Big(1 + \frac{\eta}{\lambda}\Big) \sum_k \nu_k^\star \\ &= -\frac{1}{\lambda} \mathbb{E}_{\nu^\star}[Z] - 1 - \frac{\eta}{\lambda}, \end{aligned}$$

using $\sum_k \nu_k^\star = 1$ in the last line.

Now plug this into $J(\nu^\star)$:

$$\begin{aligned} J(\nu^\star) &= \mathbb{E}_{\nu^\star}[Z] + \lambda D_{\mathrm{KL}}(\nu^\star \| \mu) \\ &= \mathbb{E}_{\nu^\star}[Z] + \lambda\Big(-\frac{1}{\lambda}\mathbb{E}_{\nu^\star}[Z] - 1 - \frac{\eta}{\lambda}\Big) \\ &= \mathbb{E}_{\nu^\star}[Z] - \mathbb{E}_{\nu^\star}[Z] - \lambda - \eta \\ &= -\lambda - \eta. \end{aligned}$$

It remains to express $-\lambda - \eta$ in terms of μ and Z. Using the expression for C above,

$$C = \exp\Big(-1 - \frac{\eta}{\lambda}\Big) = \frac{1}{\sum_j \mu_j \exp(-Z_j/\lambda)}. \tag{37}$$

Taking logs gives

$$-1 - \frac{\eta}{\lambda} = -\log \sum_j \mu_j \exp(-Z_j/\lambda), \tag{38}$$

$$\implies \quad \eta = \lambda\Big(-1 + \log \sum_j \mu_j \exp(-Z_j/\lambda)\Big). \tag{39}$$

Therefore

$$\begin{aligned} J(\nu^\star) &= -\lambda - \eta \\ &= -\lambda - \lambda\Big(-1 + \log \sum_j \mu_j \exp(-Z_j/\lambda)\Big) \\ &= -\lambda \log \sum_j \mu_j \exp(-Z_j/\lambda). \end{aligned}$$

Recognizing the sum as an expectation under μ,

$$\sum_j \mu_j \exp(-Z_j/\lambda) = \mathbb{E}_\mu\big[e^{-Z/\lambda}\big],$$

we obtain

$$J(\nu^\star) = -\lambda \log \mathbb{E}_\mu\big[e^{-Z/\lambda}\big].$$

This is exactly (34), completing the proof. □

Remark. In a general (possibly continuous) measurable space, the same derivation goes through with sums replaced by integrals and ν_k, μ_k replaced by densities (Radon–Nikodym derivatives). Equation (34) is then the standard Donsker–Varadhan variational representation.