King's Research Portal

Document Version
Peer reviewed version

Link to publication record in King's Research Portal

Bayesian Optimization with Formal Safety Guarantees via Online Conformal Prediction

Yunchuan Zhang, *Graduate Student Member, IEEE*, Sangwoo Park, *Member, IEEE*, and Osvaldo Simeone, *Fellow, IEEE*

arXiv:2306.17815v3 [cs.LG] 4 Jul 2024

Abstract—Black-box zero-th order optimization is a central primitive for applications in fields as diverse as finance, physics, and engineering. In a common formulation of this problem, a designer sequentially attempts candidate solutions, receiving noisy feedback on the value of each attempt from the system. In this paper, we study scenarios in which feedback is also provided on the *safety* of the attempted solution, and the optimizer is constrained to limit the number of unsafe solutions that are tried throughout the optimization process. Focusing on methods based on Bayesian optimization (BO), prior art has introduced an optimization scheme – referred to as SAFEOPT – that is guaranteed not to select *any* unsafe solution with a controllable probability over feedback noise as long as strict assumptions on the safety constraint function are met. In this paper, a novel BO-based approach is introduced that satisfies safety requirements irrespective of properties of the constraint function. This strong theoretical guarantee is obtained at the cost of allowing for an arbitrary, controllable but non-zero, rate of violation of the safety constraint. The proposed method, referred to as SAFE-BOCP, builds on online conformal prediction (CP) and is specialized to the cases in which feedback on the safety constraint is either noiseless or noisy. Experimental results on synthetic and real-world data validate the advantages and flexibility of the proposed SAFE-BOCP.

Index Terms—Bayesian optimization, online conformal prediction, safe exploration.

I. INTRODUCTION

A. Context and Scope

PROBLEMS as diverse as stock portfolio optimization and asset management [1], capacity allocation in energy systems [2], material discovery [3], calibration and optimization of quantum systems [4], and scheduling and optimization of wireless systems [5], [6] can all be formulated as *black-box zero-th order* optimizations. In such problems, the objective to be optimized can only be accessed on individual candidate solutions, and no further information is retrieved apart from the value of the objective. As illustrated in Fig. 1, in a common formulation of this problem, a designer sequentially attempts candidate solutions, receiving noisy feedback on the value of each attempt from the system. In this paper, we study

Figure 1. This paper studies black-box zero-th order optimization with safety constraints. At each step $t = 1, 2, ...$ of the sequential optimization process, the optimizer selects a candidate solution $\mathbf{x}_t$ and receives noisy feedback on the values of the objective function $f(\mathbf{x}_t)$ and of the constraint function $q(\mathbf{x}_t)$. Candidate solutions $\mathbf{x}_t$ yielding a negative value for the constraint function, $q(\mathbf{x}_t) < 0$, are deemed to be unsafe. We wish to keep the safety violation rate, i.e., the fraction of unsafe solutions attempted during the optimization process, below some tolerated threshold.

scenarios in which feedback is also provided on the *safety* of the attempted solution, and the optimizer is constrained to limit the number of unsafe solutions that are tried throughout the optimization process [7]–[11].

As an example, consider the problem of discovering pharmaceuticals for a particular condition (see, e.g., [12]). A pharmaceutical company may try different molecules by carrying out costly trials with patients. Such trials would return not only an indication of the effectiveness of the candidate cure, but also an indication of possible side effects. A reasonable goal is that of finding a maximally effective compound, while minimizing the number of molecules that are found to have potential side effects during the optimization process.

Typical tools for the solution of black-box zero-th order optimization construct surrogates of the objective function that are updated as information is collected by the optimizer. This can be done using tools from reinforcement learning, such as bandit optimization [13], or Bayesian optimization (BO) [14]–[17].

Focusing on methods based on BO, prior art has introduced

an optimization scheme – referred to as SAFEOPT [7], [8] – that is guaranteed not to select *any* unsafe solution with a controllable probability with respect to feedback noise. This theoretical guarantee is, however, only valid if the optimizer has access to information about the constraint function. In particular, reference [7], [8] assumes that the constraint function belongs to a reproducible kernel Hilbert space (RKHS), and that it has a known finite RKHS norm. In practice, specifying such information may be difficult, since the constraint function is a priori unknown.

In this paper, a novel BO-based approach is introduced that satisfies safety requirements irrespective of properties of the constraint function. This guarantee is obtained at the cost of allowing for an arbitrary, controllable but non-zero, rate of violation of the safety constraint. The proposed method, referred to as SAFE-BOCP, builds on online conformal prediction (CP) [18], [19], and is specialized to the cases in which feedback on the safety constraint is either noiseless or noisy.

B. Related Work

Existing constrained sequential black-box zero-th order optimizers that leverage BO, collectively referred as Safe-BO schemes, target a strict safety requirement whereby no safety violations are allowed. Accordingly, all candidate solutions attempted by the optimizer must be safe [7]–[11]. As mentioned in the previous subsection, such stringent safety requirements can only be guaranteed by making strong assumptions on the knowledge available regarding the safety constraint function. In particular, all the existing works on Safe-BO, with a notable exception of [11], either assume knowledge of the smoothness properties of the constraint function when dealing with deterministic constraint function [7]–[10], or treating the constraint function as a random realization of a Gaussian process with a known kernel when dealing with random constraint function [8].

When the mentioned assumptions or the surrogate model on the constraint function are invalid or infeasible, the existing methods cannot provide any formal safety guarantees. In order to mitigate this problem, reference [11] proposed to apply meta-learning [20] to estimate a suitable surrogate model for the constraint function using additional data that are assumed to be available from other, similar, optimization problems. However, no formal safety guarantees are available for the approach.

In a related line of work, the constrained BO approaches in [21]–[23] target a constrained optimization problem, but allow unlimited safety violations during the optimization process. More recent references [24], [25] considered an explicit budget of safety violations for probabilistic constraints, but did not provide any formal safety guarantees.

CP is a general framework for the calibration of statistical models [26]. CP methods can be applied to pre-trained machine learning models with the goal of ensuring that the model's outputs provide reliable estimates of their uncertainty. There are two main classes of CP techniques: *offline CP*, which leverages offline calibration data for this purpose [26], [27]; and *online CP*, which uses feedback on the reliability of

Table I
STATE OF THE ART ON SAFE-BO AGAINST THE PROPOSED SAFE-BOCP

	Safe-BO [7]–[11]	SAFE-BOCP (ours)
Target safety violation rate	0	(0,1)
Assumption-free safety guarantee	✗	✓

past decisions to adjust the post-processing of model's outputs [18], [19]. In both cases, CP offers theoretical guarantees on the quality of the uncertainty quantification provided by the decisions of the system.

The relevance of *online CP* for the problem of interest, illustrated in Fig. 1, is that, as the optimizer attempts multiple solutions over time, it needs to maintain an estimate of the constraint function. In order to ensure the safety of the candidate solutions selected by the optimizers, it is important that such estimates come with well-calibrated uncertainty intervals. In this paper, we leverage the theoretical guarantees of online CP in order to define novel BO-based safe optimization strategies.

The only existing combination of CP and BO we are aware of are provided by [28], which apply *offline* CP to BO for the solution of an *unconstrained* optimization problem. The approach aims at improving the acquisition function while accounting for observation noise that goes beyond the standard homoscedastic Gaussian assumption. These prior works do not address safety requirements.

C. Main Contributions

In this paper, we introduce SAFE-BOCP, a novel BO-based optimization strategy for constrained black-box zero-th order problems with safety constraints. SAFE-BOCP provides *assumptions-free* guarantees on the safety level of the attempted candidate solutions, while enabling any non-zero target safety violation level. As summarized in Table I, this contrasts with the state-of-the-art papers [7]–[11] that only target the most stringent safety constraint with no safety violations throughout the optimization process, while relying on strong assumptions on the constraint function [7]–[10].

To summarize, the main contributions of the paper are as follows:
• We introduce the deterministic SAFE-BOCP (D-SAFE-BOCP) algorithm, which assumes noiseless feedback on the constraint function and targets a flexible safety constraint on the average number of candidate solutions that are found to be unsafe. The approach is based on a novel combination of online CP and Safe-BO methods.
• For the case in which feedback on the constraint function is noisy, we introduce the probabilistic SAFE-BOCP (P-SAFE-BOCP) algorithm, which targets a flexible safety constraint on the *probability* that the average number of candidate solutions that are found to be unsafe exceeds a controllable threshold. The method relies on a "caution-increasing" back-off mechanism that compensates for the uncertainty on the safety feedback received from the system.
• We prove that both D-SAFE-BOCP and P-SAFE-BOCP meet their target safety requirements irrespective of the properties of the constraint function.

• We validate the performance of all the proposed methods and theorems on a synthetic data set and on real-world applications.

The rest of the paper is organized as follows. Sec. II formulates the constrained black-box zero-th order problem with safety constraints. The general framework of Safe-BO, as well as the representative, state-of-the-art, algorithm SAFEOPT, are reviewed in Sec. III and Sec. IV, respectively. The proposed SAFE-BOCP methods are introduced in the following sections, with D-SAFE-BOCP presented in Sec. V and P-SAFE-BOCP described in Sec. VI. Experimental results on synthetic dataset are provided in Sec. VII, and Sec. VIII demonstrates results on real-world applications. Finally, Sec. IX concludes the paper.

II. PROBLEM FORMULATION

In this section, we describe the constrained black-box zero-th order optimization problems for safety-critical scenarios studied in this work. Then, we introduce the general solution framework of interest in the next section, which is referred to as Safe-BO [7]–[11].

A. Optimization Problem and Safety Constraint

We focus on constrained optimization problems of the form

$$\max_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \quad \text{s.t.} \quad q(\mathbf{x}) \geq 0, \tag{1}$$

where objective function $f(\mathbf{x})$ and constraint function $q(\mathbf{x})$ are real valued; and $\mathcal{X}$ is some specified subset of the d-dimensional vector space $\mathbb{R}^d$. Let f^{opt} denote the maximum value of the problem (1), which we assume to be finite. We also assume that the set of optimal solutions, achieving the optimal value f^{opt}, is not empty. We write any optimal solution as $\mathbf{x}^{\text{opt}} \in \mathcal{X}$ with $f^{\text{opt}} = f(\mathbf{x}^{\text{opt}})$. Furthermore, we assume that there is a known, non-empty, set $\mathcal{S}_0 \subset \mathcal{X}$ of safe solutions, i.e.,

$$\mathcal{S}_0 \subseteq \{\mathbf{x} \in \mathcal{X} \colon q(\mathbf{x}) \geq 0\}. \tag{2}$$

This subset may be as small as a single safe solution $\mathbf{x}_0$ with $q(\mathbf{x}_0) \geq 0$, i.e., $\mathcal{S}_0 = \{\mathbf{x}_0\}$.

We address the optimization problem (1) under the following conditions.

• *Zero-th-order black-box access*: The real-valued objective function $f(\mathbf{x})$ and constraint function $q(\mathbf{x})$ are a priori unknown, and only accessible as zero-th-order black boxes. This implies that, given a candidate solution $\mathbf{x}$, the optimizer can evaluate both functions, obtaining the respective values $f(\mathbf{x})$ and $q(\mathbf{x})$. In practice, the evaluations are often noisy, resulting in the observation of noisy values $\tilde{f}(\mathbf{x})$ and $\tilde{q}(\mathbf{x})$. No other information, such as gradients, is obtained by the optimizer about the functions.

• *Efficient optimization*: The optimizer wishes to minimize the number of accesses to both functions $f(\mathbf{x})$ and $q(\mathbf{x})$, while producing a feasible and close-to-optimal solution $\mathbf{x}^* \in \mathcal{X}$. That is, we wish for the optimizer to output a vector $\mathbf{x}^* \in \mathcal{X}$ that satisfies the constraint $q(\mathbf{x}^*) \geq 0$, with an objective value

$f(\mathbf{x}^*)$ close to the maximum value f^{opt}. The performance of the optimizer can be measured by the *optimality ratio*

$$\Delta f(\mathbf{x}^*) = \frac{f(\mathbf{x}^*)}{f^{\text{opt}}}. \tag{3}$$

• *Safety*: Interpreting the inequality $q(\mathbf{x}) \geq 0$ as a safety constraint, we consider choices of the optimization variable $\mathbf{x} \in \mathcal{X}$ that result in a negative value of the constraint function $q(\mathbf{x})$ to be *unsafe*, unless the number of such violations of the constraint are kept below a threshold. Accordingly, we will require that the number of evaluations of the constraint function $q(\mathbf{x})$ that result in a violation of the inequality $q(\mathbf{x}) \geq 0$ to be no larger than a pre-determined value. We will formalize this constraint next by describing the general operation of the optimizer.

B. Sequential Surrogate-Based Safe Optimization

Starting from a given solution $\mathbf{x}_0 \in \mathcal{S}_0$ (2), the optimizer sequentially produces *candidate solutions* $\mathbf{x}_1, ..., \mathbf{x}_T \in \mathcal{X}$ across T *trials* or *iterations*. At each iteration t, the optimizer receives noisy observations of the objective value $f(\mathbf{x}_t)$ as

$$y_t = f(\mathbf{x}_t) + \epsilon_{f,t}, \tag{4}$$

as well as a noisy observation of the constraint value $q(\mathbf{x}_t)$ as

$$z_t = q(\mathbf{x}_t) + \epsilon_{q,t}, \tag{5}$$

where the observation noise for the objective, $\epsilon_{f,t} \sim \mathcal{N}(0, \sigma_f^2)$, is Gaussian with variance σ_f^2, while the observation noise for the constraint, $\epsilon_{q,t}$, can follow any distribution provided that it has a known upper bound on the one-sided right-tail probability (see Assumption 1 in Sec. VI-A for details).

We focus on optimizers that maintain *surrogate models* of functions $f(\mathbf{x})$ and $q(\mathbf{x})$ in order to select the next iterate. To elaborate, let us write as $\mathcal{O}_t$ the overall history of past iterates $(\mathbf{x}_0, ..., \mathbf{x}_t)$ and past observations $(y_0, z_0, ..., y_t, z_t)$ at the end of the t-th iteration, i.e.,

$$\mathcal{O}_t = (\mathbf{x}_0, ..., \mathbf{x}_t, y_0, ..., y_t, z_0, ..., z_t). \tag{6}$$

As we detail in the next section, the optimizer maintains probability distributions $p(f|\mathcal{O}_t)$ and $p(q|\mathcal{O}_t)$ on the functions $f(\mathbf{x})$ and $q(\mathbf{x})$ across all values $\mathbf{x} \in \mathcal{X}$ based on the available information $\mathcal{O}_t$. The distributions $p(f|\mathcal{O}_t)$ and $p(q|\mathcal{O}_t)$ summarize the belief of the optimizer regarding the values of the two functions.

At the next iteration $t + 1$, the optimizer leverages the distributions $p(f|\mathcal{O}_t)$ and $p(q|\mathcal{O}_t)$ to obtain iterate $\mathbf{x}_{t+1}$ as follows.

• *Safe set*: Using distribution $p(q|\mathcal{O}_t)$, the optimizer identifies a safe set $\mathcal{S}_{t+1} \subseteq \mathcal{X}$, containing solutions $\mathbf{x} \in \mathcal{X}$ deemed by the optimizer to be safe, i.e., to satisfy the constraint $q(\mathbf{x}) \geq 0$.

• *Acquisition*: Using distributions $p(f|\mathcal{O}_t)$ and $p(q|\mathcal{O}_t)$, the optimizer selects the next iterate $\mathbf{x}_{t+1} \in \mathcal{S}_{t+1}$, with the aim of maximizing the likelihood of obtaining a large, i.e., close to 1, optimality ratio (3).

C. Safety Constraints

We now formalize the safety constraint by distinguishing the cases in which the observations (5) of constraint function $q(\mathbf{x})$ are: (i) *noiseless*, i.e., we have $z_t = q(\mathbf{x}_t)$ in (5) with noise power $\sigma_q^2 = 0$; and (ii) *noisy*, i.e., we have a positive observation noise power $\sigma_q^2 > 0$ in (5).

1) Deterministic Safety Constraint: Noiseless observations of the constraint function values allow the optimizer to keep track of the number of iterates $\mathbf{x}_t$ that result in violations of the non-negativity constraint in problem (1). Accordingly, with $\sigma_q^2 = 0$, we impose that the non-negativity constraint $q(\mathbf{x}_t) \geq 0$ be violated no more than a tolerated fraction $\alpha \in [0, 1]$ of the T iterations. Specifically, given a *target violation rate* $\alpha \in [0, 1]$, this results in the deterministic safety requirement

$$\text{violation-rate}(T) := \frac{1}{T} \sum_{t=1}^{T} \mathbb{1}(q(\mathbf{x}_t) < 0) \leq \alpha, \qquad (7)$$

where $\mathbb{1}(\cdot)$ is the indicator function, i.e., we have $\mathbb{1}(\text{true}) = 1$ and $\mathbb{1}(\text{false}) = 0$. Therefore, in this first case, we target the maximization of function $f(x)$ subject to the safety constraint (7) on the optimization process.

2) Probabilistic Safety Constraint: In the presence of observation noise on the constraint, i.e., with a positive observation noise power $\sigma_q^2 > 0$, the optimizer cannot guarantee the deterministic constraint (7). Rather, targeting problem (1), the optimizer can only aim at ensuring that the constraint (7) be satisfied with a probability no smaller than a *target reliability level* $1 - \delta$, with $\delta \in (0, 1]$. This results in the *probabilistic safety constraint*

$$\Pr(\text{violation-rate}(T) \leq \alpha) \geq 1 - \delta, \qquad (8)$$

in which the probability is taken with respect to the observation noise variables $\{\epsilon_{q,t}\}_{t=1}^{T}$ for the constraint function $q(\mathbf{x})$ in (5). Therefore, in this second case, we target the maximization of function $f(x)$ subject to the safety constraint (8) on the optimization process.

III. SAFE BAYESIAN OPTIMIZATION

We adopt BO as the underlying surrogate-based optimization strategy. When deployed to address the problem of safe black-box optimization defined in the previous section, BO-based schemes are referred to collectively as *Safe-BO* [7]–[11]. As illustrated in Fig. 2, Safe-BO models objective function $f(\mathbf{x})$ and constraint function $q(\mathbf{x})$ by using independent Gaussian processes (GPs) as surrogate models, producing the distributions $p(f|\mathcal{O}_t)$ and $p(q|\mathcal{O}_t)$ introduced in Sec. II-B. In this section, we first review background material on GPs in Sec. III-A. Then, we discuss a general approach to define safe sets $\mathcal{S}_{t+1}$ on the basis of the current distribution $p(q|\mathcal{O}_t)$ in Sec. III-B.

A. Gaussian Process

Consider an unknown scalar-valued function $g(\mathbf{x})$ with input $\mathbf{x} \in \mathbb{R}^d$. GP models such a function by assuming that, for any collection $(\mathbf{x}_1, ..., \mathbf{x}_N)$ of inputs, the corresponding
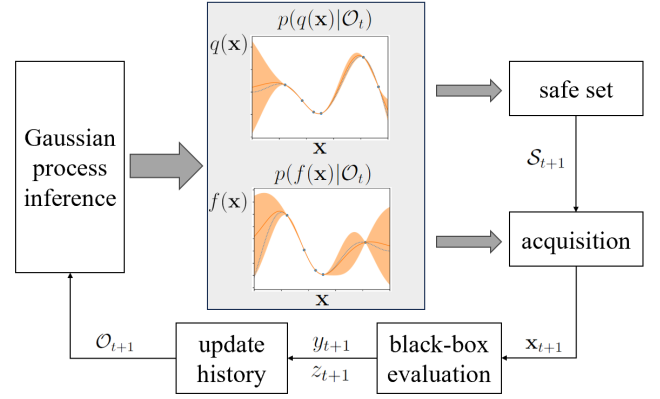

Figure 2. Block diagram of Safe-BO schemes consisting of the main steps of safe set creation, producing the safe set $\mathcal{S}_{t+1}$, and of acquisition, selecting the next iterate $\mathbf{x}_{t+1}$.

outputs $(g(\mathbf{x}_1), ..., g(\mathbf{x}_1))$ follow a multivariate Gaussian distribution. The Gaussian distribution is characterized by a mean function $\mu(\mathbf{x})$ with $\mathbf{x} \in \mathbb{R}^d$, and kernel function $\kappa(\mathbf{x}, \mathbf{x}')$ for $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^d$ [29]. An examples of a kernel function is the radial basis function (RBF) kernel

$$\kappa(\mathbf{x}, \mathbf{x}') = \exp(-h||\mathbf{x} - \mathbf{x}'||^2), \qquad (9)$$

which depends on a bandwidth parameter $h > 0$. Specifically, for given inputs $(\mathbf{x}_1, ..., \mathbf{x}_N)$, collectively denoted as $\mathbf{X}$, the output vector $(g(\mathbf{x}_1), ..., g(\mathbf{x}_1))$ follows a Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}(\mathbf{X}), \mathbf{K}(\mathbf{X}))$, with $N \times 1$ mean vector $\boldsymbol{\mu}(\mathbf{X}) = [\mu(\mathbf{x}_1), ..., \mu(\mathbf{x}_N)]^\mathsf{T}$, and $N \times N$ covariance matrix $\mathbf{K}(\mathbf{X})$ with each (n, n')-th entry given by $\kappa(\mathbf{x}_n, \mathbf{x}_{n'})$.

Assume that the output $g(\mathbf{x})$ is observed in the presence of independent Gaussian noise as

$$y = g(\mathbf{x}) + \epsilon, \qquad (10)$$

with $\epsilon \sim \mathcal{N}(0, \sigma^2)$. We write as $\mathbf{y} = [y_1, ..., y_N]^\mathsf{T}$ the $N \times 1$ vector collecting the noisy outputs (10) for inputs $(\mathbf{x}_1, ..., \mathbf{x}_N)$. An important property of GPs is that, given the history $\mathcal{O} = (\mathbf{X}, \mathbf{y})$ of previous observations $\mathbf{y}$ for inputs $\mathbf{X}$, the posterior distribution $p(y|\mathbf{x}, \mathcal{O})$ of a new output y corresponding to any input $\mathbf{x}$ has a Gaussian distribution with mean $\mu(\mathbf{x}|\mathcal{O})$ and variance $\sigma^2(\mathbf{x}|\mathcal{O})$, i.e.,

$$p(g(\mathbf{x})|\mathcal{O}) = \mathcal{N}(\mu(\mathbf{x}|\mathcal{O}), \sigma^2(\mathbf{x}|\mathcal{O})), \qquad (11a)$$

with $\quad \mu(\mathbf{x}|\mathcal{O}) = \mu(\mathbf{x}) + \boldsymbol{\kappa}(\mathbf{x})^\mathsf{T}(\mathbf{K}(\mathbf{X}) + \sigma^2\mathbf{I}_N)^{-1}(\mathbf{y} - \boldsymbol{\mu}(\mathbf{X})),$
$$(11b)$$

and $\quad \sigma^2(\mathbf{x}|\mathcal{O}) = \kappa(\mathbf{x}, \mathbf{x}) - \boldsymbol{\kappa}(\mathbf{x})^\mathsf{T}(\mathbf{K}(\mathbf{X}) + \sigma^2\mathbf{I}_N)^{-1}\boldsymbol{\kappa}(\mathbf{x}),$
$$(11c)$$

with $N \times 1$ covariance vector $\boldsymbol{\kappa}(\mathbf{x}) = [\kappa(\mathbf{x}, \mathbf{x}_1), ..., \kappa(\mathbf{x}, \mathbf{x}_N)]^\mathsf{T}$ and identity matrix $\mathbf{I}_N \in \mathbb{R}^{N \times N}$.

B. Credible Intervals and Safe Set

Let us return to the operation of sequential optimizers based on BO. As explained in the previous section, at the end of iteration t, the optimizer has attempted solutions $(\mathbf{x}_1, ..., \mathbf{x}_t)$, which are collectively referred to as $\mathbf{X}_t$. For these inputs, it has observed the noisy values $\mathbf{y}_t = [y_1, ..., y_t]^\mathsf{T}$ in (4) of the

objective function, as well as the noisy values $\mathbf{z}_t = [z_1, ..., z_t]^\mathsf{T}$ in (5) for the constraint function. As we reviewed in Sec. III-A, GPs allow the evaluation of the posterior distributions

$$p(f(\mathbf{x})|\mathcal{O}_t) = p(f(\mathbf{x})|\mathbf{X}_t, \mathbf{y}_t) \qquad (12)$$

and

$$p(q(\mathbf{x})|\mathcal{O}_t) = p(q(\mathbf{x})|\mathbf{X}_t, \mathbf{z}_t) \qquad (13)$$

for a new candidate solution $\mathbf{x}$, given the history $\mathcal{O}_t = (\mathbf{X}_t, \mathbf{y}_t, \mathbf{z}_t)$ consisted of the previous attempts $\mathbf{X}_t$ and its corresponding noisy observations $\mathbf{y}_t$ and $\mathbf{z}_t$. As we discuss next, these posterior distributions are used by Safe-BO methods to construct *credible intervals*, which quantify the residual uncertainty on the values of functions $f(\mathbf{x})$ and $q(\mathbf{x})$ at any candidate solution $\mathbf{x}$.

Introducing a *scaling parameter* $\beta_{t+1} > 0$, the credible interval for the value of the objective function $f(\mathbf{x})$ for input $\mathbf{x}$ at the end of iteration t, or equivalently at the beginning of iteration $t+1$, is defined by lower bound $f_l(\mathbf{x}|\mathcal{O}_t)$ and upper bound $f_u(\mathbf{x}|\mathcal{O}_t)$ given by

$$
\begin{aligned}
\mathcal{I}_f(\mathbf{x}|\mathcal{O}_t) &= [f_l(\mathbf{x}|\mathcal{O}_t), f_u(\mathbf{x}|\mathcal{O}_t)] \\
&= [\mu_f(\mathbf{x}|\mathbf{X}_t, \mathbf{y}_t) - \beta_{t+1}\sigma_f(\mathbf{x}|\mathbf{X}_t, \mathbf{y}_t), \\
&\qquad \mu_f(\mathbf{x}|\mathbf{X}_t, \mathbf{y}_t) + \beta_{t+1}\sigma_f(\mathbf{x}|\mathbf{X}_t, \mathbf{y}_t)],
\end{aligned} \qquad (14)
$$

where the mean $\mu_f(\mathbf{x}|\mathbf{X}_t, \mathbf{y}_t)$ and the standard deviation $\sigma_f(\mathbf{x}|\mathbf{X}_t, \mathbf{z}_t)$ are defined as in (11b) and (11c), respectively. In a similar manner, the credible interval for the constraint function $q(\mathbf{x})$ is defined as

$$
\begin{aligned}
\mathcal{I}_q(\mathbf{x}|\mathcal{O}_t) &= [q_l(\mathbf{x}|\mathcal{O}_t), q_u(\mathbf{x}|\mathcal{O}_t)] \\
&= [\mu_q(\mathbf{x}|\mathbf{X}_t, \mathbf{z}_t) - \beta_{t+1}\sigma_q(\mathbf{x}|\mathbf{X}_t, \mathbf{z}_t), \\
&\qquad \mu_q(\mathbf{x}|\mathbf{X}_t, \mathbf{z}_t) + \beta_{t+1}\sigma_q(\mathbf{x}|\mathbf{X}_t, \mathbf{z}_t)],
\end{aligned} \qquad (15)
$$

where the mean $\mu_q(\mathbf{x}|\mathbf{X}_t, \mathbf{y}_t)$ and the standard deviation $\sigma_q(\mathbf{x}|\mathbf{X}_t, \mathbf{z}_t)$ are also defined as in (11b) and (11c), respectively.

Under the Gaussian model assumed by GP, the intervals (14) and (15) include the true function values $f(\mathbf{x})$ and $q(\mathbf{x})$ for a given input $\mathbf{x}$ with probability

$$\mathrm{P}(\beta_{t+1}) = 2F(\beta_{t+1}) - 1, \qquad (16)$$

where $F(\cdot)$ is the cumulative distribution function (CDF) of standard Gaussian random variable $F(z) = \Pr(Z \le z)$ with $Z \sim \mathcal{N}(0,1)$. Therefore, the lower bounds $f_l(\mathbf{x}|\mathcal{O}_t)$ and $q_l(\mathbf{x}|\mathcal{O}_t)$ in the credible intervals (14) and (15), respectively, serve as *pessimistic* estimates of the objective and constraint values at the confidence level defined by probability $\mathrm{P}(\beta_{t+1})$. Furthermore, under the same confidence level, the upper bounds $f_u(\mathbf{x}|\mathcal{O}_t)$ and $q_u(\mathbf{x}|\mathcal{O}_t)$ in (14) and (15) describe *optimistic* estimates of the objective and constraint values, respectively. That said, it is important to stress that, since the Gaussian model assumed by GP is generally *misspecified*, there is no guarantee on the actual probability that the credible intervals $\mathcal{I}_f(\mathbf{x}|\mathcal{O}_t)$ and $\mathcal{I}_q(\mathbf{x}|\mathcal{O}_t)$ include the true values $f(\mathbf{x})$ and $q(\mathbf{x})$. These intervals, in fact, are guaranteed to include the true functions values with probability $\mathrm{P}(\beta_{t+1})$ only under the GP model.

In order to meet the safety requirement (7) or (8), Safe-BO methods define a *safe set* of candidate solutions $\mathbf{x} \in \mathcal{X}$ that are likely to satisfy the constraint $q(\mathbf{x}) \ge 0$. To this end, the optimizer selects the scaling factor β_{t+1} so as to ensure some desired "safety" probability $\mathrm{P}(\beta_{t+1})$. Then, leveraging the GP model, Safe-BO methods adopt the pessimistic estimate of the value of constraint function given by $q_l(\mathbf{x}|\mathcal{O}_t)$ in (15) as a conservative estimate of the constraint function. Accordingly, the safe set $\mathcal{S}_{t+1}$ is defined as the set of all feasible solutions $\mathbf{x} \in \mathcal{X}$ for which the conservative estimate $q_l(\mathbf{x}|\mathcal{O}_t)$ of constraint function $q(\mathbf{x})$ predicts the solution $\mathbf{x}$ to be safe, i.e.,

$$\mathcal{S}_{t+1} = \mathcal{S}(\mathcal{O}_t|\beta_{t+1}) = \{\mathbf{x} \in \mathcal{X} : q_l(\mathbf{x}|\mathcal{O}_t) \ge 0\} \cup \mathcal{S}_0. \quad (17)$$

The safe set includes the known initial set $\mathcal{S}_0$ of safe solutions in (2), ensuring a non-empty safe set [8].

Safe-BO schemes choose as the first solution $\mathbf{x}_0$ a point randomly selected from the initial safe set $\mathcal{S}_0$. For the following iterations, while all Safe-BO schemes adopt the same definition of the safe set (17), the realization of the acquisition process selecting the next iterate $\mathbf{x}_{t+1}$ differentiates the schemes proposed in prior [7]–[11]. In the next section, we specifically describe the operation of SAFEOPT [7], [8].

IV. SAFEOPT

In this section, we review SAFEOPT [7], [8] a representative state-of-the-art Safe-BO method, which will serve as a reference for the proposed SAFE-BOCP strategies introduced in the next section.

A. Scope and Working Assumptions

SAFEOPT addresses problem (1) under a strict version of the probabilistic safety constraint (8) with target violation rate $\alpha = 0$ and arbitrary target reliability level $1 - \delta$. In order to allow for a zero violation rate ($\alpha = 0$) to be a feasible goal, SAFEOPT makes the assumption that the constraint function $q(\mathbf{x})$ in (1) lies in the RKHS $\mathcal{H}_\kappa$ associated with the same kernel function $\kappa(\mathbf{x}, \mathbf{x}')$ assumed by GP inference (see Sec. III-A). In this sense, the model adopted by GP is assumed by SAFEOPT to be well specified.

Formally, the mentioned assumption made by SAFEOPT enforces that the function can be expressed as

$$q(\mathbf{x}) = \sum_{i=1}^{m} a_i \kappa(\mathbf{x}, \mathbf{x}_i) \qquad (18)$$

for some vectors $\{\mathbf{x}_i \in \mathbb{R}^d\}_{i=1}^m$, real coefficients $\{a_i\}_{i=1}^m$, and integer m. For a function $q(\mathbf{x})$ of the form (18), the *squared RKHS norm* is defined as

$$||q||_\kappa^2 = \sum_{i=1}^{m} \sum_{j=1}^{m} a_i a_j \kappa(\mathbf{x}_i, \mathbf{x}_j). \qquad (19)$$

Furthermore, a useful property of constraint function $q(\mathbf{x})$ in RKHS $\mathcal{H}_\kappa$ is that it is upper bounded by a function of their squared RKHS norm as

$$|q(\mathbf{x})| \le \kappa(\mathbf{x}, \mathbf{x})^{1/2} ||q||_\kappa \qquad (20)$$

for all values $\mathbf{x}$ in their domain. The property (20) is leveraged by SAFEOPT by assuming that the RKHS norm of the constraint function $q(\mathbf{x})$ is upper bounded by a known constant B, i.e.,

$$||q||_\kappa \leq B. \tag{21}$$

B. Safe Set Creation

Safe-BO determines the safe set $\mathcal{S}_{t+1}$ in (17) using the scaling parameter

$$\beta_{t+1} = B + 4\sigma_q\sqrt{\gamma_t + 1 - \ln(\delta)}, \tag{22}$$

where B is the constant appearing in the assumed upper bound (21); σ_q^2 is the known observation noise power in (5); $1-\delta$ is the target reliability level in (8); and γ_t is the *maximal mutual information* between the true values $(q(\mathbf{x}_1), ..., q(\mathbf{x}_t))$ of the constraint function and the corresponding t noisy observations $(\mathbf{z}_1, ..., \mathbf{z}_t)$ when evaluated under the model assumed by GP. This quantity can be evaluated as [8]

$$\gamma_t = \max_{\mathbf{X}'_t = (\mathbf{x}'_1, ..., \mathbf{x}'_t)} \left(\frac{1}{2} \log \left| \mathbf{I}_t + \sigma_q^{-2}\mathbf{K}_q(\mathbf{X}'_t) \right| \right), \tag{23}$$

where $\mathbf{I}_t$ is the $t \times t$ identity matrix and $\mathbf{K}_q(\mathbf{X}'_t)$ is the $t \times t$ covariance matrix defined in Sec. III-A. Evaluating (23) requires a maximization over all possible inputs sequences $\mathbf{X}'_t = (\mathbf{x}'_1, ..., \mathbf{x}'_t)$, hence in practice it is often addressed via greedy algorithms (see, e.g., [30]). We also observe that, in the limit of no observation noise, i.e., as $\sigma_q \to 0$, the scaling parameter (22) tends to $\beta_t = B$.

By choosing the scaling parameter β_{t+1} as in (22), under the key assumption (21), all the decisions in the safe set $\mathcal{S}_{t+1}$ (17) can be proved to be safe with high probability [8, Lemma 1] (see also [30, Theorem 6]).

C. Acquisition Process

In this section, we detail the acquisition process adopted by SAFEOPT to select the next iterate $\mathbf{x}_{t+1}$ within the safe set $\mathcal{S}_{t+1}$.

To start, SAFEOPT defines the set of *potential optimizers* $\mathcal{M}_{t+1}$ as the set of all possible solutions $\mathbf{x} \in \mathcal{S}_{t+1}$ that may increase the objective function. It also maintains a set of *possible expanders* $\mathcal{G}_{t+1}$ as the set of safe solutions that can potentially increase the size of the safe set $\mathcal{S}_{t+1}$ if selected. Then, given the potential optimizers $\mathcal{M}_{t+1}$ and the possible expanders $\mathcal{G}_{t+1}$, SAFEOPT chooses the solution $\mathbf{x} \in \mathcal{M}_{t+1} \cup \mathcal{G}_{t+1}$ that maximally reduces the larger uncertainty implied by the credible intervals (14) and (15), i.e.,

$$\mathbf{x}_{t+1} = \underset{\mathbf{x} \in \mathcal{M}_{t+1} \cup \mathcal{G}_{t+1}}{\arg \max} \max\{\sigma_f(\mathbf{x}|\mathcal{O}_t), \sigma_q(\mathbf{x}|\mathcal{O}_t)\}. \tag{24}$$

We now describe the construction of sets $\mathcal{M}_{t+1}$ and $\mathcal{G}_{t+1}$. For the first, let us recall that the lower bound $f_l(\mathbf{x}|\mathcal{O}_t)$ in the credible interval (14) can be viewed as a pessimistic estimate of the objective $f(\mathbf{x})$, while the upper bound $f_u(\mathbf{x}|\mathcal{O}_t)$ can be interpreted as an optimistic estimate of the same value. The set of potential optimizers, $\mathcal{M}_{t+1}$, includes all safe solutions $\mathbf{x} \in \mathcal{S}_{t+1}$ for which the optimistic estimate $f_u(\mathbf{x}|\mathcal{O}_t)$ is larger than

---

Algorithm 1: SAFEOPT

Input: GP priors $(\mu_f(\mathbf{x}), \kappa_f(\mathbf{x}, \mathbf{x}'))$ and $(\mu_q(\mathbf{x}), \kappa_q(\mathbf{x}, \mathbf{x}'))$, initial safe set $\mathcal{S}_0$, initial observation $\mathcal{O}_0$, assumed RKHS norm bound B, total number of optimization iterations T

Output: Decision $\mathbf{x}^*$

Initialize scaling parameters $\{\beta_t\}_{t=1}^{T+1}$ using (22), $\mathbf{x}_1 = \text{SAFEOPT}(\mathcal{O}_0|\beta_1)$

for $t = 1, ..., T$ **do**
> Observe y_t and z_t from candidate solution $\mathbf{x}_t$
> Update the observation history
> $\mathcal{O}_t = \mathcal{O}_{t-1} \cup \{\mathbf{x}_t, y_t, z_t\}$
> Update GPs with $\mathcal{O}_t$ as in (12) and (13)
> $\mathbf{x}_{t+1} = \text{SAFEOPT}(\mathcal{O}_t|\beta_{t+1})$

end

Return final decision $\mathbf{x}^* = \arg\max_{\mathbf{x} \in \mathcal{S}_{T+1}} f_l(\mathbf{x}|\mathcal{O}_T)$

SAFEOPT$(\mathcal{O}_t|\beta_{t+1})$:
> Create credible intervals $\mathcal{I}_f(\mathbf{x}|\mathcal{O}_t)$ and $\mathcal{I}_q(\mathbf{x}|\mathcal{O}_t)$ using β_{t+1} as in (14) and (15)
> Obtain safe set $\mathcal{S}_{t+1}$ as in (17)
> Update the set of potential optimizers $\mathcal{M}_{t+1}$ as in (25)
> Update the set of possible expanders $\mathcal{G}_{t+1}$ as in (27)
> **Return** the next iterate $\mathbf{x}_{t+1}$ in accordance to (24)

---

the best pessimistic estimate $f_l(\mathbf{x}|\mathcal{O}_t)$ for all safe solutions $\mathbf{x} \in \mathcal{S}_{t+1}$. This set can be expressed mathematically as

$$\mathcal{M}_{t+1} = \left\{ \mathbf{x} \in \mathcal{S}_{t+1} \Big| f_u(\mathbf{x}|\mathcal{O}_t) \geq \max_{\mathbf{x}' \in \mathcal{S}_{t+1}} f_l(\mathbf{x}'|\mathcal{O}_t) \right\}. \tag{25}$$

Note that this set is non-empty, since it includes at least the solution $\mathbf{x}$ that maximizes the lower bound $f_l(\mathbf{x}|\mathcal{O}_t)$.

The set $\mathcal{M}_{t+1}$ accounts only for the objective value to select solutions from the safe set $\mathcal{S}_{t+1}$. In contrast, the set of possible expanders considers the potential impact of a selected candidate solution on the safe set. To formalize this concept, let us write $\mathcal{S}_{t+2}(\mathbf{x})$ for the safe set (17) evaluated by extending the current history $\mathcal{O}_t$ with the pair $(\mathbf{x}, q_u(\mathbf{x}|\mathcal{O}_t))$ of candidate solution $\mathbf{x}$ and corresponding hypothetical observation of the optimistic value $q_u(\mathbf{x}|\mathcal{O}_t)$ of the constraint $q(\mathbf{x})$. Accordingly, we have

$$\mathcal{S}_{t+2}(\mathbf{x}) = \mathcal{S}\Big(\mathcal{O}_t \cup (\mathbf{x}, q_u(\mathbf{x}|\mathcal{O}_t)) \Big| \beta_{t+1}\Big), \tag{26}$$

and the set of possible expanders is defined as

$$\mathcal{G}_{t+1} = \{\mathbf{x} \in \mathcal{S}_{t+1} : |\mathcal{S}_{t+2}(\mathbf{x}) \setminus \mathcal{S}_{t+1}| > 0\}, \tag{27}$$

that is, as the set of all safe solutions that can potentially increase the size of the safe set.

After T trials, the final decision $\mathbf{x}^*$ is obtained by maximizing the pessimistic estimate $f_l(\mathbf{x}|\mathcal{O}_T)$ of the objective function that is available after the last iteration over the safe set $\mathcal{S}_{T+1}$, i.e.,

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{S}_{T+1}} f_l(\mathbf{x}|\mathcal{O}_T). \tag{28}$$

The overall procedure of SAFEOPT is summarized in Algorithm 1.

D. Safety Property

SAFEOPT was shown in [7], [8] to achieve the probabilistic safety constraint (8) with $\alpha = 0$, as long as the assumptions that the true constraint function $q(\mathbf{x})$ is of the form (18) and that the RKHS norm bound (21) holds.

Theorem 1. *(Safety Guarantee of SAFEOPT [8]) Assume that the RKHS norm of the true constraint function $q(\mathbf{x})$ is bounded by $B > 0$ as in (21). By choosing the scaling parameter β_{t+1} as in (22), SAFEOPT satisfies the probabilistic safety constraint (8) with $\alpha = 0$. Furthermore, with ideal observations of the constraint function $q(\mathbf{x})$, i.e., $\sigma_q = 0$, by choosing the scaling parameter as $\beta_{t+1} = B$, SAFEOPT meets the deterministic requirement (7) with $\alpha = 0$.*

From Theorem 1, as long as the Gaussian model assumed by GP is well specified – in the sense indicated by the RKHS form (18) with known norm upper bound B in (22) – SAFEOPT ensures safe optimization with a zero target violation rate $\alpha = 0$. In practice, however, it is hard to set a value for the constant B. Therefore, for any fixed constant B, the resulting algorithm does not have formal guarantees in terms of safety [11].

V. DETERMINISTIC SAFE-BO VIA ONLINE CONFORMAL PREDICTION

As we have reviewed in Sec. IV, in order to achieve a zero target violation rate $\alpha = 0$ in the safety constraints (7) and (8), SAFEOPT assumes that the constraint function $q(\mathbf{x})$ belongs to a specific family of functions. Other Safe-BO methods [8]–[10] also require the same assumption to guarantee the safety constraint (see Sec. I). In the following two sections, we will introduce SAFE-BOCP, a novel Safe-BO scheme that achieves the safety constraint requirements (7) or (8) without requiring *any* assumptions on the underlying constraint function $q(\mathbf{x})$. This goal is met at the cost of obtaining a non-zero, controllable, target violation rate $\alpha \in (0, 1]$ in the deterministic safety requirement (7) and in the probabilistic safety requirement (8). This section focuses on the case in which observations (5) of the constraint function are ideal, i.e., $\epsilon_{q,t} = 0$, hence aiming at achieving the deterministic safety constraint (7). The next section addresses the case with noisy observations on the constraint function.

A. Adaptive Scaling via Noiseless Feedback on Safety

As detailed in Sec. III, SAFEOPT fixes *a priori* the scaling parameters $\beta_1, ..., \beta_T$ to be used when forming the safe set (17), along with the set of potential optimizers (25) and possible expanders (27), irrespective of the actual history $\mathcal{O}_t$ of past iterates $\mathbf{X}_t$ and observations $\mathbf{y}_t$ and $\mathbf{z}_t$. This is done by leveraging the mentioned assumptions on the constraint function (18)–(21). In contrast, not relying on any assumption on the constraint function $q(\mathbf{x})$, the proposed SAFE-BOCP selects the scaling parameter β_{t+1} adaptively based on the history $\mathcal{O}_t$ by leveraging ideas from online CP [18], [19].

In order to support the adaptive selection of a scaling parameter β_{t+1} that ensures the deterministic safety constraint (7), SAFE-BOCP maintains an *excess violation rate* variable $\Delta\alpha_{t+1}$ across the iterations $t = 1, ..., T$. The variable $\Delta\alpha_{t+1}$
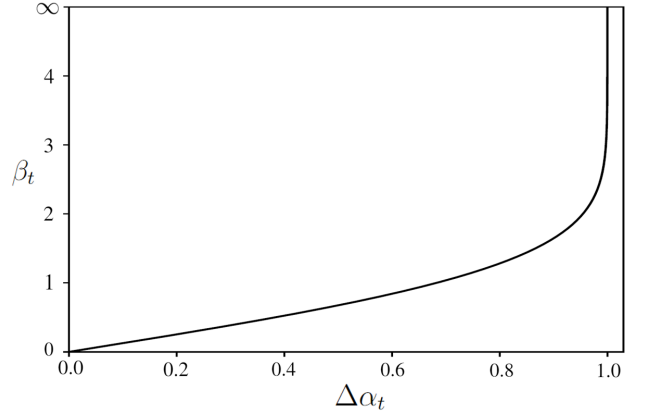

Figure 3. Function $\beta_t = \varphi(\Delta\alpha_t)$ in (34), which determines the scaling factor β_t as a function of the excess violation rate $\Delta\alpha_t$.

compares the number of previous unsafe candidate solutions $\mathbf{x}'_t$ with $t' = 1, ..., t$ to a tolerable number that depends on the target violation rate α. The main idea is to use the excess violation rate $\Delta\alpha_{t+1}$ to update the parameter β_{t+1}: A larger excess violation rate $\Delta\alpha_{t+1}$ calls for a larger value of β_{t+1} so as to ensure a more pronounced level of pessimism in the evaluation of the safe set (17). This forces the acquisition function (24) to be more conservative, driving down the excess violation rate towards a desired non-positive value.

B. D-SAFE-BOCP

To define the excess violation rate, we first introduce the *safety error signal*

$$\mathrm{err}_t = \mathbb{1}(z_t < 0), \tag{29}$$

which yields $\mathrm{err}_t = 1$ if the last iterate $\mathbf{x}_t$ was found to be unsafe based on the observation $z_t = q(\mathbf{x}_t)$, and $\mathrm{err}_t = 0$ otherwise. An important property of schemes, like SAFEOPT and D-SAFE-BOCP, that rely on the use of safe sets of the form (17) is that one can ensure a zero error signal $\mathrm{err}_t = 0$ by setting $\beta_t = \infty$. In fact, with this maximally cautious selection, the safe set $\mathcal{S}_t$ includes only the initial safe set $\mathcal{S}_0$ in (2), which consists exclusively of safe solutions.

The excess violation rate $\Delta\alpha_{t+1}$ measures the extent to which the average number of errors made so far, $1/t \cdot \sum_{t'=1}^{t} \mathrm{err}_{t'}$, exceeds an algorithmic target level α_{algo}, which will be specified later. Accordingly, the excess violation rate is updated as

$$\Delta\alpha_{t+1} = \Delta\alpha_t + \eta(\mathrm{err}_t - \alpha_{\mathrm{algo}}), \tag{30}$$

for a given update rate $\eta > 0$ and for any initialization $\Delta\alpha_1 < 1$. The relation between excess violation rate and the average number of errors becomes apparent by rewriting (30) as

$$\Delta\alpha_{t+1} = \Delta\alpha_1 + \eta \cdot \left(\sum_{t'=1}^{t} \mathrm{err}_{t'} - \alpha_{\mathrm{algo}} \cdot t \right)$$

$$= \Delta\alpha_1 + \eta \cdot t \cdot \left(\text{violation-rate}(t) - \alpha_{\mathrm{algo}} \right), \tag{31}$$

which is a linear function of the difference between the violation rate up to time t and the algorithmic target α_{algo}.

Algorithm 2: D-SAFE-BOCP

Input: GP priors $(\mu_f(\mathbf{x}), \kappa_f(\mathbf{x}, \mathbf{x}'))$ and
  $(\mu_q(\mathbf{x}), \kappa_q(\mathbf{x}, \mathbf{x}'))$, initial safe set $\mathcal{S}_0$, initial
  observation $\mathcal{O}_0$, total number of optimization
  iterations T, target violation rate α, update rate
  $\eta > 0$, initial excess violation rate $\Delta\alpha_1 < 1$.
Output: Decision $\mathbf{x}^*$
  Initialize $\mathbf{x}_1 = \text{SAFEOPT}(\mathcal{O}_0|\beta_1)$ using $\beta_1 = \varphi(\Delta\alpha_1)$
  (34), algorithmic target level α_{algo} as in (35)
  for $t = 1, ... T$ **do**
  | Observe y_t and z_t from candidate solution $\mathbf{x}_t$
  | Update the observation history
  | $\mathcal{O}_t = \mathcal{O}_{t-1} \cup \{\mathbf{x}_t, y_t, z_t\}$
  | Update GPs with $\mathcal{O}_t$ as in (12) and (13)
  | Evaluate error signal $\text{err}_t = \mathbb{1}(z_t < 0)$ as in (29)
  | Update excess violation rate
  | $\Delta\alpha_{t+1} = \Delta\alpha_t + \eta(\text{err}_t - \alpha_{\text{algo}})$ as in (30)
  | Update scaling parameter $\beta_{t+1} = \varphi(\Delta\alpha_{t+1})$ using
  | (34)
  | $\mathbf{x}_{t+1} = \text{SAFEOPT}(\mathcal{O}_t|\beta_{t+1})$ from Algorithm 1
  end
  Return final decision $\mathbf{x}^* = \arg\max_{\mathbf{x} \in \mathcal{S}_{T+1}} f_l(\mathbf{x}|\mathcal{O}_T)$

This implies that the desired safety requirement (7) can be equivalently imposed via the inequality

$$\text{violation-rate}(T) = \frac{\Delta\alpha_{T+1} - \Delta\alpha_1}{T\eta} + \alpha_{\text{algo}} \leq \alpha. \quad (32)$$

Therefore, controlling the violation rate requires us to make sure that the excess violation rate $\Delta\alpha_t$ does not grow too quickly with the iteration index t.

Intuitively, as mentioned, in order to control the value of the excess violation rate $\Delta\alpha_t$, we need to select values of β_t that increase with $\Delta\alpha_t$. To this end, as summarized in Algorithm 2, inspired by the approach introduced by [19] in the context of online CP, the proposed D-SAFE-BOCP sets the parameter β_t as

$$\beta_t = \varphi(\Delta\alpha_t), \quad (33)$$

where we have defined function

$$\varphi(\Delta\alpha_t) = F^{-1}((\text{clip}(\Delta\alpha_t) + 1)/2), \quad (34)$$

with $F^{-1}(\cdot)$ being the inverse of the function $F(\cdot)$ (16), i.e., the inverse CDF of standard Gaussian distribution, and $\text{clip}(\Delta\alpha_t) = \max\{\min\{\Delta\alpha_t, 1\}, 0\}$ being the clipping function. An illustration of the function (34) can be found in Fig. 3. Furthermore, we set the algorithmic target level as

$$\alpha_{\text{algo}} = \frac{1}{T-1}\left(T\alpha - 1 - \frac{1}{\eta} + \frac{\Delta\alpha_1}{\eta}\right). \quad (35)$$

The overall procedure of D-SAFE-BOCP is summarized in Algorithm 2. We next prove that D-SAFE-BOCP meets the reliability requirement (32).

C. Safety Guarantees

D-SAFE-BOCP is guaranteed to meet the deterministic safety constraint (7) (or equivalently (32)), as summarized in the next theorem.

Theorem 2 (Safety Guarantee of D-SAFE-BOCP)**.** *Under noiseless observations of the constraint function ($\sigma_q^2 = 0$), D-SAFE-BOCP satisfies the deterministic safety constraint (7) for any pre-determined target violation rate $\alpha \in (0, 1]$.*

Proof. Function (33) implements the following mechanism: When $\Delta\alpha_t \geq 1$, it returns $\beta_t = \infty$, i.e.,

$$\Delta\alpha_t \geq 1 \Rightarrow \beta_t = \infty. \quad (36)$$

As discussed earlier in this section, this ensures a zero error signal $\text{err}_t = 0$. With this mechanism in place, one can guarantee the upper bound

$$\Delta\alpha_{t+1} < 1 + \eta(1 - \alpha_{\text{algo}}) \quad (37)$$

for all $t \geq 1$ given the mentioned initialization $\Delta\alpha_1 < 1$. This is because a value $\Delta\alpha_t \geq 1$ would cause the update term in (30) to $-\eta\alpha_{\text{algo}} < 0$, and hence the maximum value is attained when $\Delta\alpha_t$ is approaching, but smaller than, 1, and an unsafe decision is made, causing an update equal to $\eta(1 - \alpha_{\text{algo}})$.

Plugging bound (37) back into (32), yields the upper bound on the violation rate

$$\text{violation-rate}(T) \leq \frac{1 + \eta(1 - \alpha_{\text{algo}}) - \Delta\alpha_1}{T\eta} + \alpha_{\text{algo}}. \quad (38)$$

Therefore, by setting (35), we finally verify that the deterministic safety requirement (32) is satisfied. $\square$

VI. PROBABILISTIC SAFE-BOCP

We now turn to the case in which the observations (5) of constraint function $q(\mathbf{x})$ are noisy. The main challenge in extending the approach proposed in the previous section is the fact that the error signal (29) is an unreliable indication of whether candidate $\mathbf{x}_t$ is safe or not due to the presence of the observation noise $\epsilon_{q,t}$. Accordingly, we start by proposing an alternative way to measure the excess violation rate.

A. P-SAFE-BOCP

To proceed, we assume that the observation noise $\epsilon_{q,t}$ in (5) has a known upper bound on the right-tail probability $\text{Pr}(\epsilon_{q,t} \geq \omega)$ for all $\omega \in \mathbb{R}$. This basic assumption is also adopted in the robust CP literature [31, Theorem 1]. In Sec. VI-C, we will illustrate how to further alleviate this assumption by assuming access to noise samples.

Assumption 1. *The constraint observation noise $\epsilon_{q,t}$, which is independent over $t = 1, ..., T$, has a known upper bound $F^+(\omega)$ on its one-sided right-tail probability, i.e.,*

$$\text{Pr}(\epsilon_{q,t} \geq \omega) \leq F^+(\omega) \quad (39)$$

for all $t = 1, ..., T$ and any $\omega \in \mathbb{R}$.

The main idea underlying the proposed P-SAFE-BOCP is to count as unsafe all solutions $\mathbf{x}_t$ for which the noisy observation

---

Algorithm 3: P-SAFE-BOCP

Input: GP priors $(\mu_f(\mathbf{x}), \kappa_f(\mathbf{x}, \mathbf{x}'))$ and
$(\mu_q(\mathbf{x}), \kappa_q(\mathbf{x}, \mathbf{x}'))$, initial safe set $\mathcal{S}_0$, initial
observation $\mathcal{O}_0$, total number of optimization
iterations T, target violation rate α, update rate
$\eta > 0$, initial excess violation rate $\Delta\alpha_1 < 1$

Output: Decision $\mathbf{x}^*$

Initialize $\mathbf{x}_1 = \text{SAFEOPT}(\mathcal{O}_0|\beta_1)$ using $\beta_1 = \varphi(\Delta\alpha_1)$
(34), algorithmic target level α_{algo} as in (35)

for $t = 1, ...T$ **do**

    Observe y_t and z_t from candidate solution $\mathbf{x}_t$
    Update the observation history
      $\mathcal{O}_t = \mathcal{O}_{t-1} \cup \{\mathbf{x}_t, y_t, z_t\}$
    Update GPs with $\mathcal{O}_t$ as in (12) and (13)
    Evaluate *cautious* error signal $\text{err}_t = \mathbb{1}(z_t < \omega_q)$ as
    in (40) with ω_q obtained from (41)
    Update excess violation rate
      $\Delta\alpha_{t+1} = \Delta\alpha_t + \eta(\text{err}_t - \alpha_{\text{algo}})$ as in (30)
    Update scaling parameter $\beta_{t+1} = \varphi(\Delta\alpha_{t+1})$ using
    (34)
    $\mathbf{x}_{t+1} = \text{SAFEOPT}(\mathcal{O}_t|\beta_{t+1})$ from Algorithm 1

end

Return final decision $\mathbf{x}^* = \arg\max_{\mathbf{x} \in \mathcal{S}_{T+1}} f_l(\mathbf{x}|\mathcal{O}_T)$

$z_t = q(\mathbf{x}_t) + \epsilon_{q,t}$ in (5) is smaller than some back-off threshold $\omega_q > 0$. Specifically, we define the safety error signal as

$$\text{err}_t = \mathbb{1}(z_t < \omega_q), \tag{40}$$

where the corresponding threshold ω_q is obtained as

$$\omega_q = \inf\{\omega \in \mathbb{R} : F^+(\omega) \le 1 - (1-\delta)^{\frac{1}{T}}\}. \tag{41}$$

The threshold ω_q increases with the target reliability level $1-\delta$ in the probabilistic safety constraint (8). In fact, a larger target reliability level calls for more caution in determining whether a given observation z_t of the constraint function is likely to indicate an unsafe solution or not.

The rationale behind the definitions (40)-(41) is formalized by the following lemma, which relates the true violation rate (7) to the estimated violation rate $\sum_{t=1}^{T} \text{err}_t/T$ using the error signal (40).

Lemma 1 (Estimated Violation Rate). *For any iterates* $\mathbf{x}_1, ..., \mathbf{x}_T$, *the true violation rate in* (7) *is upper bounded by the accumulated error signal rate in* (40) *with probability* $1-\delta$, *i.e.,*

$$\Pr\left(\text{violation-rate}(T) \le \frac{1}{T}\sum_{t=1}^{T} err_t\right) \ge (1 - F^+(\omega))^T$$
$$= 1 - \delta, \tag{42}$$

in which the probability is taken with respect to the observation noise variables $\{\epsilon_{q,t}\}_{t=1}^{T}$ *for the constraint function* $q(\mathbf{x})$ *in* (5).

Proof. When a candidate solution $\mathbf{x}_t$ is unsafe, i.e., when $q(\mathbf{x}_t) < 0$, the probability that the error signal err_t in (40) correctly reports an error, setting $\text{err}_t = 1$, is lower bounded

by $1 - F^+(\omega)$. Therefore, the probability that the true violation rate violation-rate(T) no larger than the estimated violation rate $\sum_{t=1}^{T} \text{err}_t/T = 1$ is lower bounded by the probability that all the errors correctly reported. This is, in turn, lower bounded by $(1 - F^+(\omega))^T$ by the independence of the observation noise variables $\{\epsilon_{q,t}\}_{t=1}^{T}$. $\square$

As specified in Algorithm 3, P-SAFE-BOCP follows the same steps in D-SAFE-BOCP with the caveat that the error signal (40) is used in lieu of (29). As we will prove next, the correction applied via the safety error signal (40) is sufficient to meet the probabilistic safety requirement (8).

B. Safety Guarantees

The safety guarantee of P-SAFE-BOCP is summarized in the following theorem.

Theorem 3 (Safety Guarantee of P-SAFE-BOCP). *Under noisy observations of the constraint function and Assumption 1, P-SAFE-BOCP satisfies the probabilistic safety constraint* (8) *for any pre-determined target violation rate* $\alpha \in (0, 1]$ *and target reliability level* $\delta \in (0, 1)$.

Proof. Using the same arguments as in the proof of Theorem 2, the estimated violation rate can be upper bounded with probability 1 as

$$\frac{1}{T}\sum_{t=1}^{T} \text{err}_t \le \frac{1 + \eta(1 - \alpha_{\text{algo}}) - \Delta\alpha_1}{T\eta} + \alpha_{\text{algo}}. \tag{43}$$

Using this bound with Lemma 1, we conclude that, with probability at least $1 - \delta$, in which the probability is taken over the observation noise variables $\{\epsilon_{q,t}\}_{t=1}^{T}$, we have bound on the *true violation rate*

$$\text{violation-rate}(T) \le \frac{1}{T}\sum_{t=1}^{T} \text{err}_t \le \alpha, \tag{44}$$

which recovers the probabilistic safety constraint (8). $\square$

C. Data-Driven Probability Bound

A possible challenge in applying P-SAFE-BOCP in practice is the fact that an upper bound $F^+(\omega)$ on the probability $\Pr(\epsilon_{q,t} \ge \omega)$ may not be known *a priori*. In this subsection, we provide a data-driven approach for evaluating an upper bound on the probability $\Pr(\epsilon_{q,t} \ge \omega)$, assuming only access to independent and identically distributed (i.i.d.) observation noise samples.

Lemma 2 (Estimated Upper Bound). *Assume access to i.i.d. observation noise samples* $\{\epsilon_{q,i}\}_{i=1}^{m}$. *The empirical estimate of the right-tail probability*

$$\hat{F}^+(\omega) = \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}(\epsilon_{q,i} > \omega), \tag{45}$$

when offset by $\psi > 0$, *provides an upper bound on* $\Pr(\epsilon_{q,i} > \omega)$ *with probability*

$$\Pr\left(\Pr(\epsilon_{q,i} \ge \omega, \forall \omega \in \mathbb{R}) \le \hat{F}^+(\omega) + \psi\right)$$
$$\ge 1 - \exp(-2m\psi^2) \tag{46}$$

for any $\psi > \sqrt{\ln 2/2m}$.

Lemma 2 is a direct application of Dvoretsky-Kiefer-Wolfowitz inequality [32].

Consequently, by using $\hat{F}^+(\omega) + \psi$ in lieu of $F^+(\omega)$ in (41), we have the following modified safety guarantee of P-SAFE-BOCP.

Corollary 1. *Under noisy observations of the constraint function, P-SAFE-BOCP with $\hat{F}^+(\omega) + \psi$, for any $\psi > 0$, in lieu of $F^+(\omega)$ in (41) satisfies the guarantee*

$$\Pr(\text{violation-rate}(T) \leq \alpha) \geq (1 - \exp(-2m\psi^2))(1 - \delta) \tag{47}$$

for any pre-determined target violation rate $\alpha \in (0, 1]$ and target reliability level $\delta \in (0, 1)$, where the probability is taken with respect to the observation noise variables $\{\epsilon_{q,t}\}_{t=1}^{T}$ as well as the m i.i.d. noise samples in (45).

Corollary 1 is obtained by combining Lemma 2 and Theorem 3. Intuitively, with an increasing number m of the constraint observation noise samples, the tightness of the safety guarantee in Theorem 3 is enhanced as a result of increasingly accurate observation noise estimation.

VII. NUMERICAL RESULTS FOR A SYNTHETIC BENCHMARK

In this section, we detail experimental results aimed at comparing SAFE-BOCP with SAFEOPT [8] on a synthetic benchmark inspired by [8].

A. Synthetic Dataset

In a manner similar to [8], we focus on a synthetic setting with a scalar optimization variable $\mathbf{x} \in \mathbb{R}$ in which the objective function $f(\mathbf{x})$ is a realization of a GP with zero mean and RBF kernel $\kappa^*(\mathbf{x}, \mathbf{x}')$ (9) with bandwidth $h^* = 1/1.62$, while the constraint function $q(\mathbf{x})$ is a function in this RKHS $\mathcal{H}_{\kappa^*}$ which has the form (18) with coefficients $\{a_i\}_{i=0}^{10} = [-0.05, -0.1, 0.3, -0.3, 0.5, 0.5, -0.3, 0.3, -0.1, -0.05]$ and scalars $\{\mathbf{x}_i\}_{i=1}^{10} = [-9.6, -7.4, -5.5, -3.3, -1.1, 1.1, 3.3, 5.5, 7.4, 9.6]$. Accordingly, the constraint function $q(\mathbf{x})$ has RKHS norm $||q||_{\kappa^*} = 1.69$ in (21). In order to investigate the impact of misspecification of GP (see Sec. IV-A) on Safe-BO including the proposed SAFE-BOCP, we consider the two cases: (*i*) *well-specified GP* that uses $\kappa^*(\mathbf{x}, \mathbf{x}')$ for the GP kernel, i.e., $\kappa(\mathbf{x}, \mathbf{x}') = \kappa^*(\mathbf{x}, \mathbf{x}')$; (*ii*) *misspecified GP* that uses RBF kernel with smaller bandwidth $h = 1/14.58 < h^*$, i.e., $\kappa(\mathbf{x}, \mathbf{x}') \neq \kappa^*(\mathbf{x}, \mathbf{x}')$, with unknown $||q||_{\kappa}$.

As discussed throughout the paper, the scaling parameter for the constraint function $q(\mathbf{x})$ in (15) is *a priori* determined by (22) for SAFEOPT, and is *adapted* by feedback via $\beta_{t+1} = \varphi(\Delta\alpha_{t+1})$ (30) for the proposed SAFE-BOCP, while we fix the scaling parameter for the objective function $f(\mathbf{x})$ in (14) to 3 since it does not affect the safety guarantee for both SAFEOPT (see [30, Theorem 6]) and SAFE-BOCP. The objective observation noise variance is set to $\sigma_f^2 = 2.5 \times 10^{-3}$; and the initial safe decision is chosen as $\mathbf{x}_0 = 0$ for which we have $q(\mathbf{x}_0) = 0.946 > 0$. For SAFE-BOCP, we set the
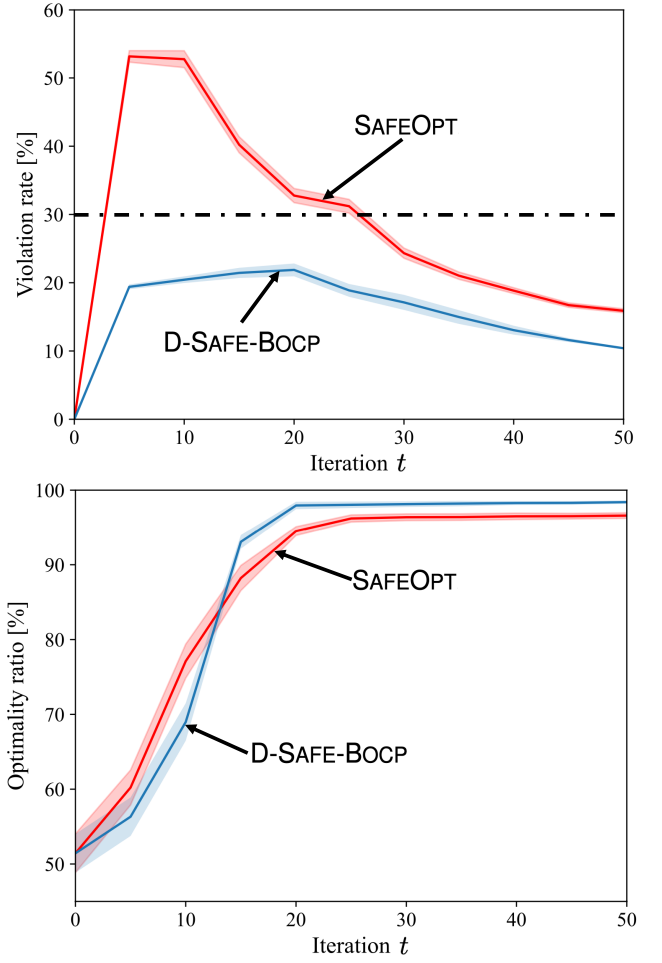

Figure 4. Violation-rate(t) (top) and optimality ratio (bottom) against iteration index t with target violation rate $\alpha = 0.3$ (dot-dashed line), update rate $\eta = 2$, misspecified kernel bandwidth $h = 1/14.58$, RKHS norm bound $B = ||q||_{\kappa^*}$ and total number of iteration $T = 50$.

update rate in (30) to $\eta = 2.0$. All results are averaged over 1,000 experiments, with error bars shown to encompass 95% of the realizations. Each experiment corresponds to a random draw of the objective function and to random realization of the observation noise signals. The implementation examples of the synthetic benchmark can be found here[1].

B. Deterministic Safety Requirement

As explained in Sec. IV, SAFEOPT requires the GP model for the constraint function $q(\mathbf{x})$ to be well specified (18)–(21) in order to meet safety conditions. To study the impact of violations of this assumption, we start by considering the noiseless case, i.e., $\sigma_q^2 = 0$, and we vary the kernel bandwidth h adopted for the GP models used as surrogates for the objective and constraint functions as discussed earlier.

Fig. 4 shows the violation rate and optimality ratio against the iteration index t. For D-SAFE-BOCP, we set the update rate as $\eta = 2$ and the target violation rate to $\alpha = 0.3$, while SAFEOPT assumes target $\alpha = 0$ with RKHS norm bound

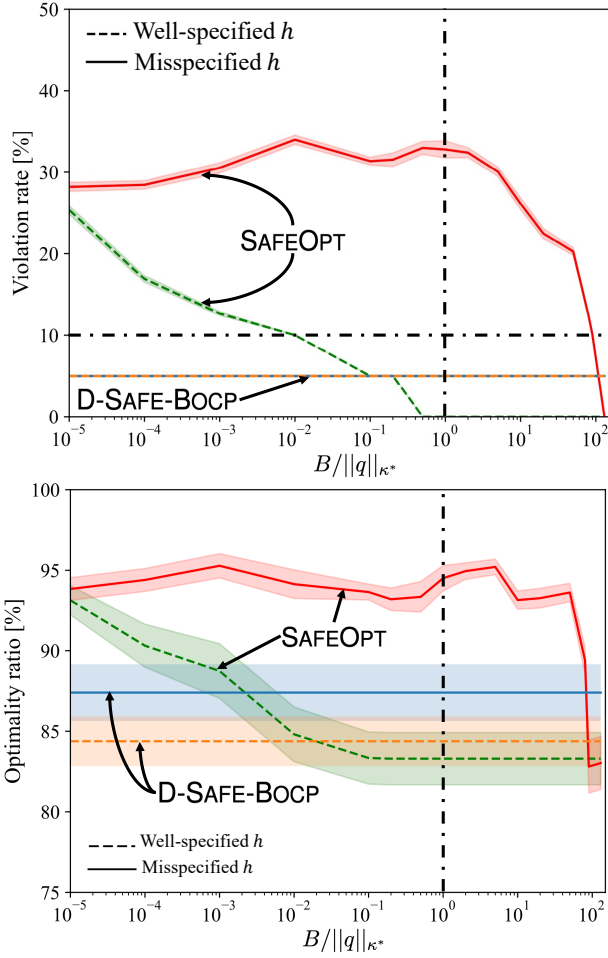[1]https://github.com/yunchuan-zhang/safe-bocp

Figure 5. Violation rate (7) (top) and optimality ratio (3) (bottom) against the ratio between the RKHS norm bound B assumed by GP and the actual norm $||q||_{\kappa^*}$ in (21). The dashed lines are obtained with well-specified GP models, which corresponds to kernel bandwidth $h = 1/1.69$ (same one for $\kappa^*(\mathbf{x}, \mathbf{x}')$), while the solid lines are obtained with misspecified GP models, having kernel bandwidth $h = 1/14.58$.

Figure 6. Violation-rate(T) and optimality ratio for different target violation rates α for D-SAFE-BOCP, with update rate $\eta = 2$, misspecified kernel bandwidth $h = 1/14.58$, and RKHS norm bound $B = ||q||_{\kappa^*}$. The background colors represent intervals in which the safety requirement (7) is met (see text for an explanation).

$B = ||q||_{\kappa^*}$. For both schemes, the total number of iterations is $T = 50$, and the misspecified GP with RBF kernel bandwidth $h = 1/14.58$ is adopted.

The violation rate obtained by SAFEOPT is above the target $\alpha = 0.3$ for a significant interval of time t, and it progressively falls below the target with a larger t, while D-SAFE-BOCP meets the deterministic safety requirement (7) with the pre-determined target $\alpha = 0.3$ across all iterations. Furthermore, the optimality ratio obtained by D-SAFE-BOCP is larger than SAFEOPT after iteration $t = 13$, converging to 97.5% at iteration $t = 20$. In contrast, SAFEOPT converges to optimality ratio of 94.5% at iteration $t = 25$, at which point the target safety level $\alpha = 0.3$ is violated.

Fig. 5 shows the violation-rate(T) in (7) with $T = 20$, as well as the optimality ratio (3), as a function of constant B assumed by SAFEOPT for both well-specified and misspecified GPs, and the target violation rate is set to $\alpha = 0.1$. Note that the performance of D-SAFE-BOCP does not depend on the value of B, which is an internal parameter for SAFEOPT, but it is affected by the choice of parameter h. By Theorem
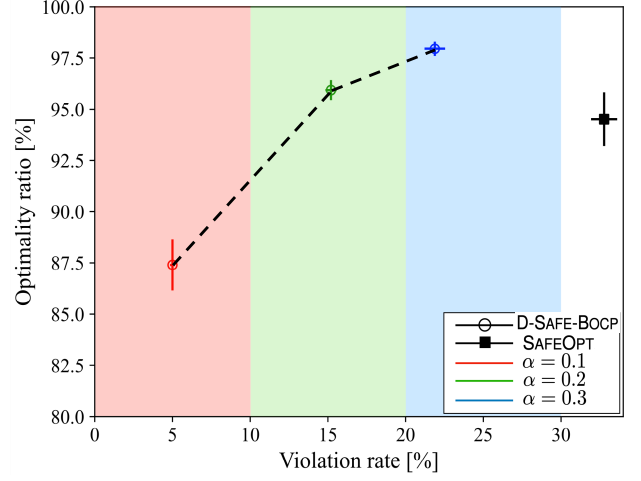
1, any value $B \geq ||q||_{\kappa}$ in (21) guarantees the safety of SAFEOPT. However, since RKHS norm for the misspecified GP is generally unknown, we plot violation rate and optimality ratio as functions of the ratio $B/||q||_{\kappa^*}$, to highlight the two regimes with well specified and misspecified value of B.

Confirming Theorem 1, with a ratio $B/||q||_{\kappa^*} \geq 1$ for the well-specified GP with kernel $\kappa(\mathbf{x}, \mathbf{x}') = \kappa^*(\mathbf{x}, \mathbf{x}')$, SAFEOPT is seen to strictly satisfy the deterministic safety constraint (7), since the violation rate is equal to zero, as per its target. Instead, when $B/||q||_{\kappa^*} < 1$, and/or when the GP is misspecified, i.e., $\kappa(\mathbf{x}, \mathbf{x}') \neq \kappa^*(\mathbf{x}, \mathbf{x}')$, the violation rate exceeds the target α. In contrast, D-SAFE-BOCP obtains a violation rate below the target α, irrespective of kernel bandwidth h assumed in GP.

In terms of optimality ratio, in the regime $B/||q||_{\kappa^*} \geq 1$, with a well-specified GP parameter h, SAFEOPT achieves around 83%, while D-SAFE-BOCP obtains the larger optimality ratio 84.5%. In contrast, with a misspecified value h, D-SAFE-BOCP achieves an optimality ratio around 87.5%, while the optimality ratio of SAFEOPT is larger, but this comes at the cost of the violation of the safety requirement. Note that a misspecified value of the kernel bandwidth h does not necessarily reduce the performance of D-SAFE-BOCP, which is improved in this example.

The trade-off between violation rate and optimality ratio is studied in Fig. 6 by varying the target violation rate α for D-SAFE-BOCP. For each value of α, we show the achieved pair of violation rate and optimality ratio, along with the corresponding realization ranges along the two axes. Recall that for SAFEOPT the assumed target is $\alpha = 0$, and hence one pair is displayed. We focus here on the misspecified GP case, i.e., $\kappa(\mathbf{x}, \mathbf{x}') \neq \kappa^*(\mathbf{x}, \mathbf{x}')$, while the SAFEOPT parameter B is selected to the "safe" value $B = ||q||_{\kappa^*}$, which is unaware of kernel misspecification.

For each value of $\alpha \in \{0.1, 0.2, 0.3\}$, the figure highlights the intervals of violation rates that meet the safety requirement
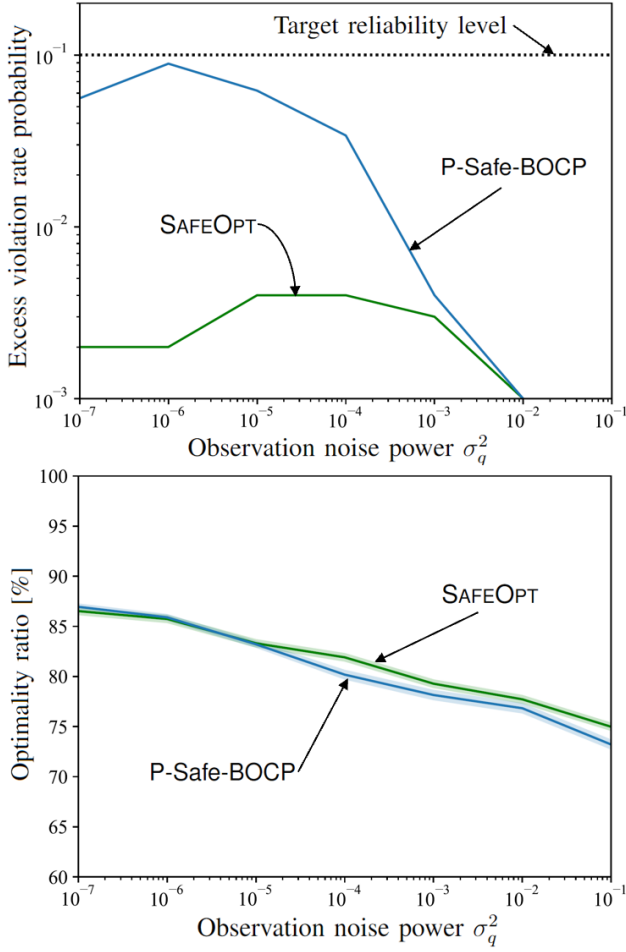
Figure 7. Probability of excessive violation rate (8) (top) and optimality ratio (3) (bottom) as a function of constraint observation noise power σ_q^2, with update rate $\eta = 2$, RKHS norm bound $B = 10||q||_{\kappa^*}$, and well-specified kernel bandwidth $h = 1/1.62$.
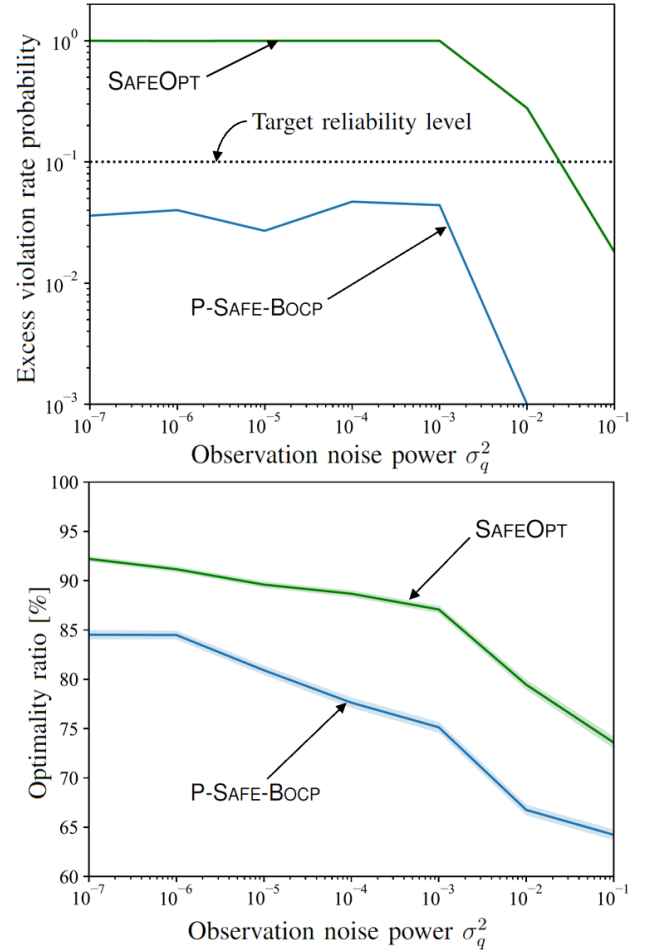

Figure 8. Excess violation rate probability (8) (top) and optimality ratio (3) (bottom) as a function of constraint observation noise power σ_q^2, with update rate $\eta = 2$, RKHS norm bound $B = 10||q||_{\kappa^*}$, and misspecified kernel bandwidth $h = 1/14.58$.

(7) using different colors. Specifically, for $\alpha = 0.1$, all violation rates below 0.1 are acceptable, as denoted by the red interval; for $\alpha = 0.2$, all violation rates in the red and green intervals are acceptable; and for $\alpha = 0.3$, all violation rates below in the cyan, green, and red interval meet the safety constraint.

The figure shows that the violation rate obtained by SAFEOPT exceeds its target $\alpha = 0$, and thus the safety requirement is violated. In contrast, as per the theory developed in this paper, D-SAFE-BOCP meets violation-rate requirement for all values of the target α. Moreover, as the tolerated violation rate α increases, the optimality ratio of D-SAFE-BOCP is enhanced, indicating a trade-off between the two metrics. When increasing the target violation rate α, D-SAFE-BOCP raises the algorithmic target level α_{algo} in (35), making it possible for the optimizer to reduce the time spent under the maximally cautious scaling $\beta_t = \infty$ in (33). Consequently, with a larger α, the optimality ratio of D-SAFE-BOCP gains from more explorations of the objective function $f(\mathbf{x})$.

C. Probabilistic Safety Constraint

We now turn to considering scenarios with Gaussian observation noise $\sigma_q^2 > 0$, and aim at evaluating the performance in terms of probabilistic safety requirement (8) and optimality ratio (3). We set the total number of iterations $T = 25$, the target reliability level $1 - \delta = 0.9$ with target violation rate $\alpha = 0.1$ for P-SAFE-BOCP, and with $\alpha = 0$ for SAFEOPT in accordance with SAFEOPT's design. For the latter scheme, we set the "safe" value $B = 10||q||_{\kappa^*}$, while we consider both the well-specified kernel bandwidth $h = h^* = 1/1.62$, and the misspecified one $h = 1/14.58 < h^*$, as considered also in the previous set of experiments. For all schemes, the excess violation rate probability in (8) is obtained by averaging over 10,000 realizations.

We plot the excess violation rate probability (8) and the optimality ratio in Fig. 7 and Fig. 8 against the observation noise power σ_q^2. The first figure corresponds to the case of a well-specified kernel bandwidth while for the second we adopted misspecified value. Confirming the theory, in the former case, both SAFEOPT and P-SAFE-BOCP attain an excess violation rate probability below the target level $1 - \delta$. In contrast, for a misspecified kernel, SAFEOPT can only satisfy

the constraint (8) for sufficiently large observation noise, but P-SAFE-BOCP still meets the probability safety constraint (8). We note that a larger observation noise is beneficial to SAFEOPT in terms of safety since it forces a larger level of pessimism in the definition of the safe set $\mathcal{S}_{t+1}$ in (17).

In terms of optimality ratio, larger observation noise power σ_q^2 generally yields a degraded optimality ratio. In the well-specified regime considered in Fig. 7, both schemes have comparable performance and the optimality ratio gap is no more than 5%. In the misspecified regime demonstrated in Fig. 8, the performance levels are not comparable, since the gains of recorded for SAFEOPT come at the cost of violations of the safety constraint (8), except for a sufficiently large observation noise power, here $\sigma_q^2 \geq 0.1$.

VIII. NUMERICAL RESULTS FOR REAL WORLD APPLICATIONS

In this section, we compare SAFEOPT [7] and SAFE-BOCP in two real-world applications, with the goal of validating the safety gains obtained by the proposed method along the optimization process.

A. Safe Movie Recommendation

As in [7], consider a system that sequentially recommends movies to a user. Each user assigns a score from 1 to 5 to a recommended movie. Following standard matrix factorization algorithms, we introduce a feature vector $\mathbf{x} \in \mathbb{R}^d$ for each movie. Accordingly, selecting a movie amounts to choosing a vector $\mathbf{x}$ within a set of possible movies. Denote as $r(\mathbf{x})$ the rating assigned by a user to movie $\mathbf{x}$. A recommendation is deemed to be unsafe if the user assigns it a rating strictly smaller than 4, i.e., if $r(\mathbf{x}) < 4$. Accordingly, we set both objective function $f(\mathbf{x})$ and constraint function $q(\mathbf{x})$ to be equal to $f(\mathbf{x}) = q(\mathbf{x}) = r(\mathbf{x}) - 4$. We focus on the deterministic safety constraint (7), since the ratings are assumed to be observed with no noise.

To define a GP model for the function that maps a movie feature vector $\mathbf{x}$ to a rating $r(\mathbf{x})$, we need to specify a kernel function, which describes the similarity between movies. As in [7], we adopt the linear kernel

$$\kappa(\mathbf{x}, \mathbf{x}') = \mathbf{x}^\mathsf{T} \mathbf{x}', \qquad (48)$$

for any two movie feature vectors $\mathbf{x}$ and $\mathbf{x}'$.

The feature vectors $\mathbf{x}$ for movies are optimized using the MovieLens-100k dataset [33], which includes sparse rating observations of 1,680 movies from 943 users. Specifically, as in [7], we randomly select 200 users to form the training data set, and we set $d = 20$ for the size of the feature vectors. Training applies the standard matrix factorization algorithm [34]. For testing, we pick the 10 test users, not selected for training, that have the most rated movies, and remove the movies with no rating from the possible selections.

Since the true underlying function that maps movie feature vector $\mathbf{x}$ to rating $r(\mathbf{x})$ is unknown, it is not possible to evaluate the RKHS norm $\|q\|_\kappa$ in (19) required by SAFEOPT. Accordingly, as in [8], we set $B = 3$ a priori for SAFEOPT. In this experiment, we run both SAFEOPT and D-SAFE-BOCP for
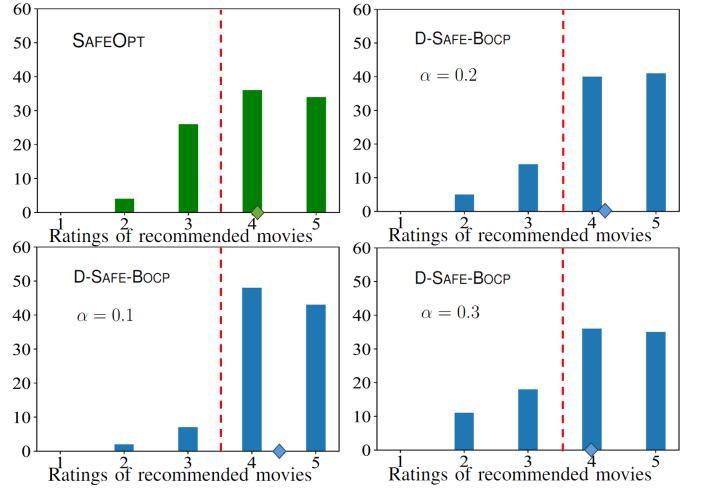

Figure 9. Histograms of the ratings of recommended movies by SAFEOPT, as well by D-SAFE-BOCP under different target violation rates α. The dashed lines represent the safety threshold for the recommendations, and the marker on the horizontal axis represents the average rating of the recommendations.

$T = 100$ iterations on the selected 10 test users. We randomly select a movie rated as 4 for each test user as the initial starting point $\mathbf{x}_0$, and set the update rate $\eta = 10$ for D-SAFE-BOCP.

To evaluate the performance of both schemes, we show in Fig. 9 the histograms of the ratings across all selected movies during the optimization procedure. The vertical dashed line represents the safety threshold between safe and unsafe recommendations. The marker on the horizontal axis marks the average rating. For D-SAFE-BOCP we have the flexibility to vary the target violation rate α, while we recall that for SAFEOPT the target is $\alpha = 0$.

The top-left panel of Fig. 9 shows that SAFEOPT does not meet the safety requirement (7) with $\alpha = 0$ owing to the mismatch between the assumptions made by the scheme and the true, unknown, constraint function. The remaining panels demonstrate that, in contrast, D-SAFE-BOCP can correctly control the fraction α of unsafe recommendations.

B. Chemical Reaction Optimization

Finally, we consider the plug flow reactor (PFR) problem introduced in [35][2], which seeks for optimal chemical reaction parameters $\mathbf{x} \in [140, 200] \times (0, 1] \subset \mathbb{R}^2$, with the first dimension being the temperature ($^\circ C$) and the second being the pH value. The goal is to maximize the yield (%), which we set as the objective $f(\mathbf{x})$, while keeping an acceptable selectivity level (%), which we denote as $s(\mathbf{x})$. We refer to [35] for a precise definition of these terms.

A reaction vector is deemed to be unsafe if the resulting selectivity level is lower than the corresponding yield, hence we define the constraint function as $q(\mathbf{x}) = s(\mathbf{x}) - f(\mathbf{x})$. We assume the presence of non-zero Gaussian observation noise z_t for the constraint function, i.e., $\sigma_q^2 > 0$. Accordingly, we focus on the probabilistic safety constraint (8), and compare

---

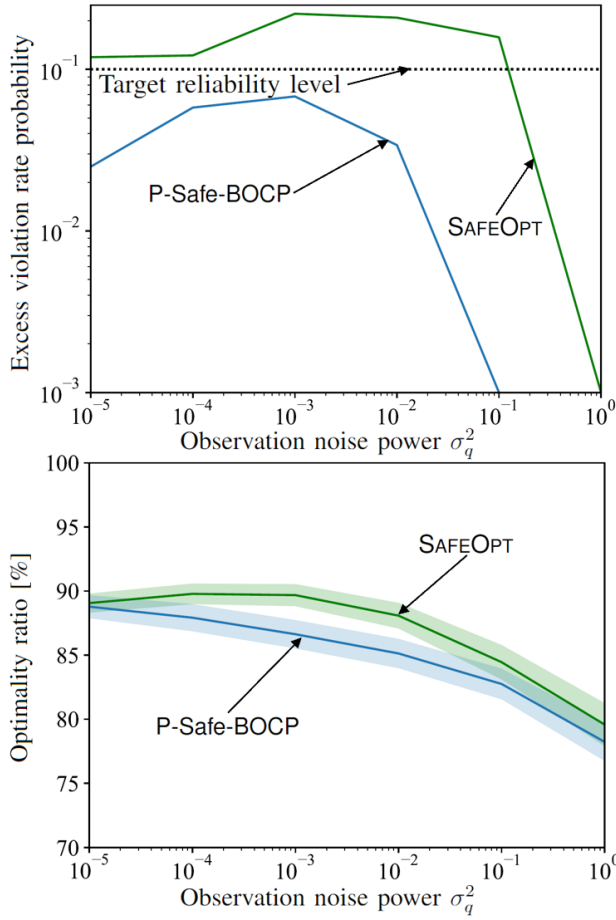[2]Simulator available at https://github.com/VlachosGroup/Fructose-HMF-Model

Figure 10. Probability of excessive violation rate (8) (top) and optimality ratio (3) (bottom) as a function of constraint observation noise power σ_q^2, with update rate $\eta = 2$, RKHS norm bound $B = 3$, and kernel bandwidth $h = 1/2.88$ for the chemical reaction problem.

the performance of SAFEOPT and P-SAFE-BOCP. We adopt GP surrogates model for both $f(\mathbf{x})$ and $q(\mathbf{x})$ with RBF kernel having bandwidth $h = 1/2.88$.

Similar to Sec. VIII-A, since the smoothness property of the true underlying functions $q(\mathbf{x})$ is unknown, we assume the constant $B = 3$ for SAFEOPT [8]. The initial decision $\mathbf{x}_0$ is randomly chosen among the a priori known safe decisions that satisfy the constraint $q(\mathbf{x_0}) \geq 0$, and we set the total number of optimization round to be $T = 50$. Other settings are as in Sec. VII-A.

In a similar manner to Sec. VII-C, we demonstrate the excess violation rate probability (8) and the optimality ratio in Fig. 10 as a function of the observation noise power σ_q^2. Confirming the discussion in Sec. VII-C and the theory, P-SAFE-BOCP is seen to meet the probabilistic safety constraint (8) irrespective of observation noise power, while SAFEOPT can only attain an excess violation rate probability below the target $1 - \delta$ when the observation noise power is sufficiently large.

IX. CONCLUSIONS

In this work, we have introduced SAFE-BOCP, a novel BO-based zero-th order sequential optimizer that provably guarantees safety requirements irrespective of the properties of the constraint function. The key mechanism underlying SAFE-BOCP adapts the level of pessimism adopted during the exploration of the search space on the basis of noisy safety feedback received by the system. From synthetic experiment to real-world applications, we have demonstrated that the proposed SAFE-BOCP performs competitively with state-of-the-art schemes in terms of optimality ratio, while providing for the first time assumption-free safety guarantees.

Although in this work we have built on SAFEOPT for the acquisition process, the proposed framework could be generalized directly to any other Safe-BO schemes, such as GOOSE [9]. Other possible extensions include accounting for multiple constraints, multi-fidelity approximations on objective and constraints [36], as well as taking into account contextual information during the optimization process [37].

REFERENCES

[1] R. O. Michaud and R. O. Michaud, *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation*. Oxford University Press, 2008.

[2] Y. Wang, R. Li, H. Dong, Y. Ma, J. Yang, F. Zhang, J. Zhu, and S. Li, "Capacity planning and optimization of business park-level integrated energy system based on investment constraints," *Energy*, vol. 189, p. 116345, 2019.

[3] S. Xu, J. Li, P. Cai, X. Liu, B. Liu, and X. Wang, "Self-improving photosensitizer discovery system via Bayesian search with first-principle simulations," *Journal of the American Chemical Society*, vol. 143, no. 47, pp. 19769–19777, 2021.

[4] C. L. Cortes, P. Lefebvre, N. Lauk, M. J. Davis, N. Sinclair, S. K. Gray, and D. Oblak, "Sample-efficient adaptive calibration of quantum networks using Bayesian optimization," *Physical Review Applied*, vol. 17, no. 3, p. 034067, 2022.

[5] W. Zhang, M. Derakhshani, G. Zheng, C. S. Chen, and S. Lambotharan, "Bayesian optimization of queuing-based multi-channel URLLC scheduling," *IEEE Transactions on Wireless Communications*, 2022.

[6] Y. Zhang, O. Simeone, S. T. Jose, L. Maggi, and A. Valcarce, "Bayesian and multi-armed contextual meta-optimization for efficient wireless radio resource management," *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 5, pp. 1282–1295, 2023.

[7] Y. Sui, A. Gotovos, J. Burdick, and A. Krause, "Safe exploration for optimization with Gaussian processes," in *Proceedings of International Conference on Machine Learning*, Lille, France, 2015.

[8] F. Berkenkamp, A. Krause, and A. P. Schoellig, "Bayesian optimization with safety constraints: Safe and automatic parameter tuning in robotics," *Machine Learning*, pp. 1–35, 2021.

[9] M. Turchetta, F. Berkenkamp, and A. Krause, "Safe exploration for interactive machine learning," in *Proceedings of Advances in Neural Information Processing Systems*, Vancouver, Canada, 2019.

[10] Y. Sui, V. Zhuang, J. Burdick, and Y. Yue, "Stagewise safe Bayesian optimization with Gaussian processes," in *Proceedings of International Conference on Machine Learning*, Stockholm, Sweden, 2018.

[11] J. Rothfuss, C. Koenig, A. Rupenyan, and A. Krause, "Meta-learning priors for safe Bayesian optimization," in *Proceedings of Conference on Robot Learning*, Atlanta, GA, USA, 2023.

[12] Y. Bengio, S. Lahlou, T. Deleu, E. J. Hu, M. Tiwari, and E. Bengio, "Gflownet foundations," *Journal of Machine Learning Research*, vol. 24, no. 210, pp. 1–55, 2023.

[13] A. Slivkins *et al.*, "Introduction to multi-armed bandits," *Foundations and Trends® in Machine Learning*, vol. 12, no. 1-2, pp. 1–286, 2019.

[14] J. Mockus, "Global optimization and the Bayesian approach," *Bayesian Approach to Global Optimization: Theory and Applications*, pp. 1–3, 1989.

[15] P. I. Frazier, "A tutorial on Bayesian optimization," *arXiv preprint arXiv:1807.02811*, 2018.

[16] L. Maggi, A. Valcarce, and J. Hoydis, "Bayesian optimization for radio resource management: Open loop power control," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 7, pp. 1858–1871, 2021.

[17] D. Eriksson, M. Pearce, J. Gardner, R. D. Turner, and M. Poloczek, "Scalable global optimization via local Bayesian optimization," in *Proceedings of Advances in Neural Information Processing Systems*, Vancouver, Canada, 2019.

[18] I. Gibbs and E. Candes, "Adaptive conformal inference under distribution shift," in *Proceedings of Advances in Neural Information Processing Systems*, Virtual, 2021.

[19] S. Feldman, L. Ringel, S. Bates, and Y. Romano, "Achieving risk control in online learning settings," *Transactions on Machine Learning Research*, 2023.

[20] T. Hospedales, A. Antoniou, P. Micaelli, and A. Storkey, "Meta-learning in neural networks: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 9, pp. 5149–5169, 2021.

[21] M. A. Gelbart, J. Snoek, and R. P. Adams, "Bayesian optimization with unknown constraints," in *Proceedings of the 30th Conference on Uncertainty in Artificial Intelligence*, Quebec, Canada, 2014.

[22] J. M. Hernández-Lobato, M. A. Gelbart, R. P. Adams, M. W. Hoffman, and Z. Ghahramani, "A general framework for constrained Bayesian optimization using information-based search," *Journal of Machine Learning Research*, vol. 17, no. 1, p. 5549–5601, 2016.

[23] J. R. Gardner, M. J. Kusner, Z. Xu, K. Q. Weinberger, and J. P. Cunningham, "Bayesian optimization with inequality constraints," in *Proceedings of International Conference on Machine Learning*, Beijing, China, 2014.

[24] A. Marco, A. von Rohr, D. Baumann, J. M. Hernández-Lobato, and S. Trimpe, "Excursion search for constrained Bayesian optimization under a limited budget of failures," *arXiv preprint arXiv:2005.07443*, 2020.

[25] C. Yu, J. Cao, and A. Rosendo, "Learning to climb: Constrained contextual Bayesian optimisation on a multi-modal legged robot," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 9881–9888, 2022.

[26] V. Vovk, A. Gammerman, and G. Shafer, *Algorithmic Learning in a Random World*, vol. 29. Springer, 2005.

[27] A. N. Angelopoulos and S. Bates, "A gentle introduction to conformal prediction and distribution-free uncertainty quantification," *Foundations and Trends® in Machine Learning*, vol. 16, no. 4, pp. 494–591, 2023.

[28] S. Stanton, W. Maddox, and A. G. Wilson, "Bayesian optimization with conformal prediction sets," in *Proceedings of International Conference on Artificial Intelligence and Statistics*, Valencia, Spain, 2023.

[29] C. E. Rasmussen, *Gaussian Processes in Machine Learning*, pp. 63–71. Springer, 2004.

[30] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Information-theoretic regret bounds for Gaussian process optimization in the bandit setting," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 3250–3265, 2012.

[31] S. Feldman, B.-S. Einbinder, S. Bates, A. N. Angelopoulos, A. Gendler, and Y. Romano, "Conformal prediction is robust to dispersive label noise," in *Proceedings of Symposium on Conformal and Probabilistic Prediction with Applications*, pp. 624–626, Limassol, Cyprus, 2023.

[32] P. Massart, "The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality," *The Annals of Probability*, pp. 1269–1283, 1990.

[33] F. Harper and J. Konstan, "The movielens datasets: History and context," *ACM Transactions on Interactive Intelligent Systems (TIIS)*, vol. 5, no. 4, 2016.

[34] D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proceedings of Advances in Neural Information Processing Systems*, Denver, USA, 2000.

[35] P.-L. Kang, C. Shang, and Z.-P. Liu, "Glucose to 5-hydroxymethylfurfural: Origin of site-selectivity resolved by machine learning based reaction sampling," *Journal of the American Chemical Society*, vol. 141, no. 51, pp. 20525–20536, 2019.

[36] Y. Zhang, S. Park, and O. Simeone, "Multi-fidelity Bayesian optimization with across-task transferable max-value entropy search," *arXiv preprint arXiv:2403.09570*, 2024.

[37] D. Widmer, D. Kang, B. Sukhija, J. Hübotter, A. Krause, and S. Coros, "Tuning legged locomotion controllers via safe Bayesian optimization," in *Proceedings of Conference on Robot Learning*, pp. 2444–2464, 2023.