Journal of Machine Learning Research 27 (2026) 1-42 Submitted 12/23; Revised 02/25; 12/25; Published 6/26

Information-Theoretic Safe Bayesian Optimization

Alessandro G. Bottero ALESSANDROGIACOMO.BOTTERO@BOSCH.COM
Bosch Corporate Research, TU Darmstadt
Robert-Bosch-Campus 1, 71272 Renningen (Germany)

Carlos E. Luis CARLOSENRIQUE.LUISGONCALVES@BOSCH.COM
Bosch Corporate Research, TU Darmstadt

Julia Vinogradska JULIA.VINOGRADSKA@BOSCH.COM
Bosch Corporate Research

Felix Berkenkamp FELIX.BERKENKAMP@BOSCH.COM
Bosch Corporate Research

Jan Peters JAN.PETERS@TU-DARMSTADT.DE
TU Darmstadt, German Research Center for AI (DFKI), Hessian.AI

Editor: Pradeep Ravikumar

Abstract

We consider a sequential decision making problem, where we aim to optimize an unknown function via noisy evaluations that do not violate an *a priori* unknown (safety) constraint. As both the objective and the constraint are unknown, a common approach is to model them using a Gaussian process and restrict evaluations to those regions that are safe with high probability. Most current methods rely on a discretization of the domain and cannot be directly extended to the continuous case. Moreover, they make regularity assumptions about the safety constraint that introduce an additional critical hyperparameter. Instead, in this paper, we propose an information-theoretic safe exploration criterion that exploits solely the GP posterior to identify the safe parameters that are most informative about the safety of other parameters. The combination of this exploration criterion with a well known Bayesian optimization acquisition function yields a novel safe Bayesian optimization selection criterion. Our approach is naturally applicable to continuous domains and does not require additional explicit hyperparameters. We theoretically analyze the method and show that the proposed acquisition function can learn about the reachable safe optimum up to arbitrary precision, while only evaluating safe parameters with high probability. Empirical evaluations demonstrate data-efficiency and scalability of our approach.

Keywords: Bayesian Optimization, Safe Exploration, Safe Bayesian Optimization, Information-Theoretic Acquisition, Gaussian Processes

1. Introduction

In many sequential decision making problems, the goal is to optimize an unknown objective function by iteratively selecting parameters to evaluate. In real-world applications such as robotics (Berkenkamp et al., 2016a), mechanical systems (Schillinger et al., 2017) or medicine (Sui et al., 2015), however, we are not free to evaluate all parameters, as these applications are often subject to additional safety constraints that we cannot violate during the exploration process (Dulac-Arnold et al., 2019). To complicate things further, in many

cases the safety constraints are *a priori* unknown as well, meaning that we do not know the set of safe parameters and cannot just perform unconstrained optimization in that set. Rather, we need to actively and carefully learn about the safety constraints without violating them. That is, in order to find the safe parameters that maximize the given objective, we also need to learn about what parameters are safe, and, most importantly, we need to do so by only evaluating parameters that are known to be safe. In this context, the well known exploration-exploitation dilemma becomes even more challenging, since one not only has to explore within the region currently classified as safe to find the optimum, but one must also promote the exploration of those areas that are informative about the safety of currently uncertain parameters, even though they might be unlikely to contain the safe optimum.

Most existing methods take a Bayesian optimization (BO) approach (Mockus, 1974; Jones et al., 1998) to tackle the safe optimization problem: they place a Gaussian process (GP) prior (Rasmussen and Williams, 2006) over both the constraint and the objective and use the GP posterior to guide exploration by restricting evaluations to those parameters that do not violate the constraint with high probability. To learn about the safety of parameters, many approaches evaluate the safe parameters with the largest posterior standard deviation, either over the whole set of safe parameters (Schreiter et al., 2015), or within a region close to the boundary of this set, defined in terms of the safety constraint's Lipschitz constant (Sui et al., 2015, 2018). However, uncertainty about the constraint and the optimum is just a proxy objective that only indirectly learns about the safety of parameters and the safe optimum. Moreover, relying on the Lipschitz constant of the safety constraint introduces an additional hyperparameter to tune. Consequently, these methods leave room to improve data efficiency and at the same time reduce the number of hyperparameters of the model.

In this paper, we tackle this challenge and develop a safe BO approach that improves on data efficiency and that does not require additional hyperparameters like the Lipschitz constant of the safety constraint. Specifically, we draw inspiration from information-theoretic acquisition functions, which have proven successful in solving the BO problem in a sample efficient manner (Hennig and Schuler, 2012; Hernández-Lobato et al., 2014; Wang and Jegelka, 2017), and develop ISE-BO, a novel information based safe BO acquisition function.

1.1 Contributions

In this paper, we introduce Information-Theoretic Safe Exploration and Optimization (ISE-BO), a novel information-theoretic acquisition function for safe Bayesian optimization. ISE-BO extends our previous work (Bottero et al., 2022)—where we considered solely the safe exploration problem—by seeking to maximize the information gain about both the safety of parameters and about the value of the safe optimum to efficiently solve the safe BO problem.

Safe Exploration As in safe BO one cannot evaluate unsafe parameters, one task of the acquisition function is to identify safe parameters from which we can infer the safety of uncertain parameters. For such safe exploration component, the acquisition function that we propose in this paper relies on the purely exploratory ISE criterion from our preliminary work (Bottero et al., 2022), which we recall, together with its derivation, in Section 3.1.

Safe Optimization While safe exploration plays a crucial role in the context of safe BO, the acquisition function must also allow us to identify the safe optimum. To this end, in this paper we combine the ISE exploration criterion with the well known max-value entropy search (MES) (Wang and Jegelka, 2017) acquisition function that aims to find the optimum by reducing the entropy of the distribution of its value. The resulting algorithm directly optimizes for information gain about both the safety of parameters and the value of the safe optimum. In this way, it achieves data-efficiency without manually restricting evaluated parameters to be in specifically designed areas (like the boundary of the safe set), and outperforms existing methods in scenarios where the posterior variance alone is not enough to identify good evaluation candidates, as in the case of heteroskedastic observation noise. Moreover, our proposed selection criterion does not require additional modeling assumptions beyond the GP posterior and is directly applicable to continuous domains.

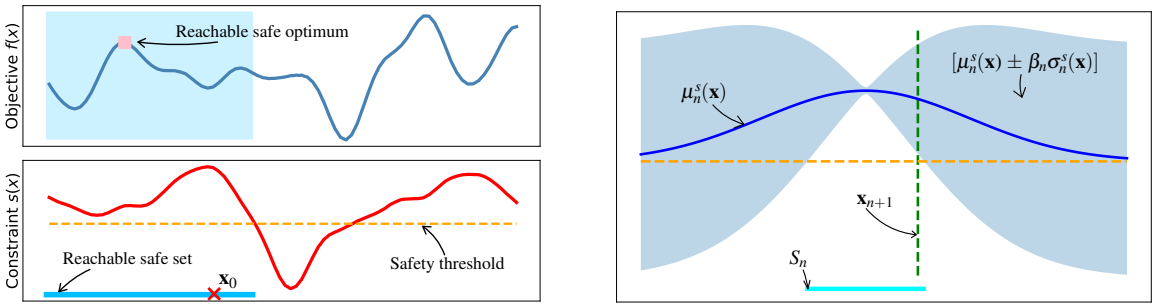
Theoretical Analysis The method that we propose offers theoretical guarantees that we show in Section 4. We first prove an exploration guarantee for the ISE component of the acquisition function: we formally define the notion of a *largest reachable safe set* as the largest set of safe parameters that we can expect a safe BO algorithm to discover, and then we show that in the discrete setting the ISE exploration criterion converges to this largest reachable safe set. Subsequently, we extend the theoretical analysis to the full ISE-BO acquisition function, showing that it learns about the reachable safe optimum to arbitrary precision. This analysis integrates and extends the theoretical results in (Bottero et al., 2022) in two major ways: first, Bottero et al. (2022) do not define a concept of convergence for the safe exploration, leaving the question open about what portion of the total set of safe parameters the algorithm can discover and classify as safe; secondly, since Bottero et al. (2022) only consider safe exploration, they do not prove any result concerning the uncertainty reduction about the safe optimum.

1.2 Related Work

Constrained BO In unconstrained BO (Mockus, 1974; Shahriari et al., 2016), the goal is to find the parameters that maximize an *a priori* unknown objective function. The presence of constraints that the final solution has to satisfy gives rise to two different problems: constrained BO, if we can violate the constraints during exploration, and safe BO, if not. In the context of constrained BO, Schonlau et al. (1998); Gramacy and Lee (2011) consider variations of the expected improvement acquisition function (Jones et al., 1998), while Gelbart et al. (2014) propose to combine typical BO acquisition functions with the probability of satisfying the constraint. Instead, Gotovos et al. (2013) propose an uncertainty-based criterion that learns about the feasible region of parameters, while Perrone et al. (2019) modify the Max-value Entropy Search selection criterion (Wang and Jegelka, 2017) to also include information about the constraint in the acquisition function. These methods exploit both safe and unsafe evaluations to find the feasible optimum.

Safe BO On the other hand, if we are not allowed to evaluate unsafe parameters, the problem becomes harder, as there is an *a priori* unknown set of unsafe parameters that we cannot evaluate, and we must instead find the safe optimum using only information coming from safe evaluations. Because of this limitation, in safe BO safe exploration becomes a necessary sub-routine that we need in order to learn about the safety of parameters. To this end, in the context of active learning, Schreiter et al. (2015) globally learn about the constraint by evaluating the most uncertain parameters, while Sui et al. (2015) were the first to cast the full safe optimization problem in terms of safe BO, and formulated SafeOpt, a safe optimization method that extends the approach by Schreiter et al. (2015) to joint exploration and optimization, and improves its efficiency by explicitly restricting safe exploration to the boundary of the safe set. Building on SafeOpt, Sui et al. (2018) further propose StageOpt, which additionally separates the exploration and optimization phases. Both of these algorithms assume access to the Lipschitz constant of the safety constraint to define parameters close to the boundary of the safe set. In practice, however, the real Lipschitz constant is not available, since the safety constraint is unknown. As a consequence, its value becomes an extra hyperparameter that one needs to tune or estimate from the noisy observations. As Sui et al. (2018) note, computing an accurate estimate can be difficult, and using the wrong one can result in either slow exploration or unsafe evaluations, depending on whether one chooses a value that is too small or too large. These methods have been extended to multiple constraints by Berkenkamp et al. (2016a), while Kirschner et al. (2019) scale them to higher dimensions with LineBO, which explores in low-dimensional sub-spaces. To improve computational costs, Duivenvoorden et al. (2017) suggest a continuous approximation to SafeOpt without providing exploration guarantees. Chan et al. (2023), on the other hand, take a more directed approach, by searching the most promising parameter also outside of the current safe set and then evaluating the most uncertain safe parameter in the neighborhood of the projection of such promising parameter on the boundary of the safe set. In the closely related context of safe active learning, Li et al. (2022) choose the parameters to evaluate as those with the highest predictive entropy and high probability of being safe.

Information-based Approaches All of the safe BO methods that we mentioned rely on function uncertainty to drive exploration. In the context of unconstrained BO, however, information-based selection criteria have often been shown to outperform other methods. In particular, Hennig and Schuler (2012); Hernández-Lobato et al. (2014); Wang and Jegelka (2017) select the parameters that maximize the information gain about the location of the optimum or the optimum value, while Fröhlich et al. (2020) consider the information under noisy parameters. Hvarfner et al. (2023), on the other hand, consider the entropy over the joint optimal probability density over both input and output spaces. We draw inspiration from these methods to define a selection criterion that directly maximizes the information gain about both the safety of parameters and the value of the safe optimum. In parallel independent work, Hübotter et al. (2024) proposed ITL, an active learning algorithm that



(a) Problem components.

(b) Safe set based on GP posterior.

Figure 1: In (a) we illustrate the safe optimization task. Based on the unknown safety constraint s, we are only allowed to evaluate parameters $\boldsymbol{x}$ with values $s(\boldsymbol{x})$ above the safety threshold. Starting from a safe seed $\boldsymbol{x}_0$, a safe optimization strategy needs to find the optimum of the unknown objective f within the largest safe region of the parameter space that is safely reachable from $\boldsymbol{x}_0$. This example also shows that, in general, the reachable safe optimum is not the global safe optimum, since safe regions could exist that are not reachable from the safe seed without evaluating uncertain or unsafe parameters. In (b) we show an example of the posterior mean and confidence interval of the GP model for s, together with the corresponding safe set S_n, as defined in (5). To guarantee safety with high probability, the next parameter that a safe BO algorithm chooses to evaluate, $\boldsymbol{x}_{n+1}$, must be within this set.

aims at learning the values of a function in a target set $\mathcal{A}$ by actively sampling observations in a sample space $\mathcal{S}$, whose relation to $\mathcal{A}$ can be arbitrary. To do so, ITL maximizes the mutual information between the prediction targets in $\mathcal{A}$ and the next observation in $\mathcal{S}$. While addressing a more general problem, when the target space is the set of potential maximizers and the sample space the set of safe parameters, ITL effectively becomes a safe BO algorithm.

Further Applications The problem of safe exploration is not unique to safe BO, but arises also in other settings. Notable examples are Markov decision processes (MDP) (Moldovan and Abbeel, 2012; Hans et al., 2008) and coverage control (Wang and Hussein, 2012). In the context of MDPs, Turchetta et al. (2016, 2019) traverse the MDP to learn about the safety of parameters using methods that, at their core, explore using the same ideas as SafeOpt and StageOpt to select parameters to evaluate, while Prajapat et al. (2022) apply these concepts to the problem of multi agent coverage control. Consequently, the information-based method that we propose to guide safe exploration is also directly applicable to their setting.

2. Problem Statement and Preliminaries

In this section, we introduce the problem and the notation that we use throughout the paper. We are given an unknown and expensive to evaluate function $f : \mathcal{X} \to \mathbb{R}$ that we aim to optimize under a safety constraint defined via another unknown function $s : \mathcal{X} \to \mathbb{R}$, s.t.

parameters that satisfy $s(\boldsymbol{x}) \geq 0$ are classified as safe, while the others as unsafe [1]

$$\begin{aligned} &\max_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x}) \\ &\text{s.t. } s(\boldsymbol{x}) \geq 0. \end{aligned} \tag{1}$$

Since both the objective f and the safety constraint s are unknown, we cannot rely on their explicit form to solve (1). Rather, we aim to find the safe optimum by sequentially evaluating both functions at parameters $\{\boldsymbol{x}_n\}_{n=0}^N \subset \mathcal{X}$. Specifically, we assume that the sequential evaluations at $\{\boldsymbol{x}_n\}_{n=0}^N$ produce noisy observations that we have access to: $y_n^f \coloneqq f(\boldsymbol{x}_n) + \nu_n^f$ and $y_n^s \coloneqq s(\boldsymbol{x}_n) + \nu_n^s$, which are corrupted by additive homoskedastic Gaussian noise $\nu_n^{f/s}$ with zero mean and standard deviation $\sigma_\nu^{f/s}$. However, unlike constrained BO, here we are not allowed to ever evaluate parameters that violate the safety constraints. That is, all the parameters $\{\boldsymbol{x}_n\}_{n=0}^N$ at which we evaluate f and s must satisfy $s(\boldsymbol{x}_n) \geq 0$ with high probability. Since before starting to evaluate s we do not have any information about what parameters are safe, in order to start exploring safely we further assume that we have access to a safe seed $\boldsymbol{x}_0$ that satisfies the safety condition, $s(\boldsymbol{x}_0) \geq 0$.

While (1) represents the ideal goal of any safe BO algorithm, without stronger assumptions about the safety constraint s, it is in general not possible to find the global optimum that (1) implies without ever evaluating unsafe parameters. That is, for some configurations of i) safe seed $\boldsymbol{x}_0$, ii) shape of safety constraint, and iii) location of global safe optimum, it is not possible to reach the global safe optimum by only evaluating parameters that are known to be safe. One example is the case of a disconnected safe set, where safe seed and global safe optimum are separated by a vast unsafe region. Therefore, we instead aim to find the safe optimum that is reachable from the given safe seed $\boldsymbol{x}_0$, i.e., the parameter that we can reach from $\boldsymbol{x}_0$ by evaluating only safe parameters and that yields the largest value of f among all such reachable parameters. We illustrate this safe optimization task in Figure 1a, where we highlight in blue the safe area reachable from $\boldsymbol{x}_0$ and the optimum of f within this region as the safe optimization goal. In this figure, we also see that the reachable safe optimum is not necessarily equal to the global safe optimum (1). In Section 2.3 we provide more intuition on this subject, together with a formal definition of the largest region that is safely reachable from $\boldsymbol{x}_0$, while in Section 4 we analyze the convergence properties of our proposed algorithm.

2.1 Probabilistic Model

As both f and s are unknown, and as we only have access to the noisy evaluations $y_n^{f/s}$, it is in general not feasible to select parameters that are safe with certainty. Instead, we provide high-probability safety guarantees by exploiting a probabilistic model for f and s. To this end, we follow Berkenkamp et al. (2016a), extend the domain to $\tilde{\mathcal{X}} \coloneqq \mathcal{X} \times \{0, 1\}$, and define the function $h : \tilde{\mathcal{X}} \to \mathbb{R}$ as

$$h(\boldsymbol{x}, i) = \begin{cases} f(\boldsymbol{x}), & \text{if } i = 0 \\ s(\boldsymbol{x}), & \text{if } i = 1. \end{cases} \tag{2}$$

Furthermore, we assume that h has bounded norm in the Reproducing Kernel Hilbert Space (RKHS) (Schölkopf and Smola, 2002) $\mathcal{H}_k$ associated to some kernel $k : \tilde{\mathcal{X}} \times \tilde{\mathcal{X}} \to \mathbb{R}$ with

1. This choice is without loss of generality, since we can incorporate any known non-zero safety threshold $\tau(\boldsymbol{x})$ in the constraint: $s'(\boldsymbol{x}) \coloneqq s(\boldsymbol{x}) - \tau(\boldsymbol{x})$.

$k\big((\boldsymbol{x}, i), (\boldsymbol{x}', j)\big) \leq 1$: $\|h\|_{\mathcal{H}_k} = B < \infty$. This assumption allows us to use a Gaussian process (GP) (Srinivas et al., 2010) as probabilistic model for h.

A Gaussian process is a stochastic process specified by a mean function $\mu : \tilde{\mathcal{X}} \to \mathbb{R}$ and a kernel $k : \tilde{\mathcal{X}} \times \tilde{\mathcal{X}} \to \mathbb{R}$ (Rasmussen and Williams, 2006). It defines a probability distribution over real-valued functions on $\tilde{\mathcal{X}}$, such that any finite collection of function values at parameters $[(\boldsymbol{x}_1, i_1), \dots, (\boldsymbol{x}_n, i_n)]$ is distributed as a multivariate normal distribution. The GP prior can then be conditioned on (noisy) function evaluations $\mathcal{D}_n = \{\big((\boldsymbol{x}_j, i_j), y_j\big)\}_{j=1}^n$. If the noise at each observation is Gaussian, $\nu_n \sim \mathcal{N}(0, \sigma_\nu^2)$ $\forall n$, then the resulting posterior is also a GP with posterior mean and variance given by

$$\begin{aligned}
\mu_n(\boldsymbol{x}, i) &= \mu(\boldsymbol{x}, i) + \boldsymbol{k}(\boldsymbol{x}, i)^\top (\boldsymbol{K} + \boldsymbol{I}\sigma_\nu^2)^{-1}(\boldsymbol{y} - \boldsymbol{\mu}), \\
\sigma_n^2(\boldsymbol{x}, i) &= k\big((\boldsymbol{x}, i), (\boldsymbol{x}, i)\big) - \boldsymbol{k}(\boldsymbol{x}, i)^\top (\boldsymbol{K} + \boldsymbol{I}\sigma_\nu^2)^{-1}\boldsymbol{k}(\boldsymbol{x}, i),
\end{aligned}$$

where $\boldsymbol{\mu} \coloneqq [\mu(\boldsymbol{x}_1, i_1), \dots \mu(\boldsymbol{x}_n, i_n)]$ is the mean vector at parameters $(\boldsymbol{x}_j, i_j) \in \mathcal{D}_n$ and $[\boldsymbol{y}]_j \coloneqq y_j$ the corresponding vector of observations. We have $\big[\boldsymbol{k}(\boldsymbol{x}, i)\big]_j \coloneqq k\big((\boldsymbol{x}, i), (\boldsymbol{x}_j, i_j)\big)$, the kernel matrix has entries $[\boldsymbol{K}]_{lm} \coloneqq k\big((\boldsymbol{x}_l, i_l), (\boldsymbol{x}_m, i_m)\big)$, and $\boldsymbol{I}$ is the identity matrix. In the following, we assume without loss of generality that the prior mean is identically zero: $\mu(\boldsymbol{x}, i) \equiv 0$.

From the GP that models h, it is straightforward to recover the posterior mean and variance for the objective f and the constraint s

$$\begin{aligned}
\mu_n^f(\boldsymbol{x}) &= \mu_n(\boldsymbol{x}, 0), \quad \mu_n^s(\boldsymbol{x}) = \mu_n(\boldsymbol{x}, 1), \\
\sigma_n^f(\boldsymbol{x}) &= \sigma_n(\boldsymbol{x}, 0), \quad \sigma_n^s(\boldsymbol{x}) = \sigma_n(\boldsymbol{x}, 1).
\end{aligned} \tag{3}$$

2.2 Safe Set

Using the previous assumptions, we can construct high-probability confidence intervals on the safety constraint values $s(\boldsymbol{x})$. Concretely, Chowdhury and Gopalan (2017) show that, for any $\delta > 0$, it is possible to find a sequence of positive numbers $\{\beta_n\}$ such that the true function value $h(\boldsymbol{x}, i)$ is within the confidence interval $\big[\mu_n(\boldsymbol{x}, i) \pm \beta_n \sigma_n(\boldsymbol{x}, i)\big]$ with probability at least $1 - \delta$, jointly for all $(\boldsymbol{x}, i) \in \tilde{\mathcal{X}}$ and $n \geq 1$. As Chowdhury and Gopalan (2017) show, a possible choice for such sequence is

$$\beta_n \coloneqq B + R\sqrt{2\big(\ln(e/\delta) + \gamma_n\big)}, \tag{4}$$

where R is the smallest positive number that satisfies both $\mathbb{P}\big\{|\nu_n^s| \geq \lambda\big\} \leq 2\exp\{-\lambda^2/R^2\}$ and $\mathbb{P}\Big\{|\nu_n^f| \geq \lambda\Big\} \leq 2\exp\{-\lambda^2/R^2\}$, while γ_n is the maximum information capacity of the chosen kernel after n iterations: $\gamma_n = \max_{|D|=n} I\big(\boldsymbol{h}(D); \boldsymbol{y}(D)\big)$ (Srinivas et al., 2010; Contal et al., 2014). From these confidence intervals, using (3) we can derive analogous intervals for s, $s(\boldsymbol{x}) \in \big[\mu_n^s(\boldsymbol{x}) \pm \beta_n \sigma_n^s(\boldsymbol{x})\big]$, which we use to define the notion of a *safe set*

$$S_n \coloneqq \{\boldsymbol{x} \in \mathcal{X} : \mu_n^s(\boldsymbol{x}) - \beta_n \sigma_n^s(\boldsymbol{x}) \geq 0\} \cup S_{n-1}; \quad S_0 \coloneqq \boldsymbol{x}_0, \tag{5}$$

namely the set that contains all parameters whose β_n-lower confidence bound is above the safety threshold, together with all parameters that were in previous safe sets. Thanks to

the set union in the definition of S_n, at each iteration the safe set either expands or remains unchanged, guaranteeing that the safe seed $\boldsymbol{x}_0$ is always contained in it. We illustrate such a safe set in Figure 1b. As a consequence of this definition, thanks to the properties of the confidence intervals $\left[\mu_n^s(\boldsymbol{x}) \pm \beta_n \sigma_n^s(\boldsymbol{x})\right]$, we also know that all parameters in S_n are safe with probability of at least $1-\delta$ jointly over all iterations n, as we show in Section 4.

2.3 Largest Reachable Safe Set and Safe Optimum

At the beginning of this section, we stated that the goal of a safe optimization algorithm is to find the optimum within the largest safe set reachable from the safe seed $\boldsymbol{x}_0$, like the one shown in Figure 1a. To make this statement precise, we now need to define what we mean with *largest reachable safe set*.

Intuition Intuitively, given that we can only evaluate parameters that are safe with high probability and that at the beginning we only know that $\boldsymbol{x}_0$ is safe, we first need a way to infer what parameters are also safe, assuming that we learn the value $s(\boldsymbol{x}_0)$ up to a precision of ε. We can then repeat this process recursively with all the new parameters that we have learned to be safe, until we cannot classify any new parameter as safe. The set of safe parameters that we obtain at the end of this process is the largest reachable safe set.

A useful tool for the theoretical analysis that allows us to formalize such intuition is a one-step *expansion operator* R_ε, such that, given a set $\Omega \subset \mathcal{X}$, $R_\varepsilon(\Omega)$ consists of all the parameters that we can assume safe once we know $s(\boldsymbol{x})$ up to ε for all $\boldsymbol{x}$ in Ω. With such operator at hand, we can define the largest reachable safe set $S_\varepsilon(\boldsymbol{x}_0)$ as the set obtained by recursively applying R_ε to $\boldsymbol{x}_0$ an infinite number of times $S_\varepsilon(\boldsymbol{x}_0) \coloneqq \lim_{n\to\infty} R_\varepsilon^n(\boldsymbol{x}_0)$.

Depending on the setting and on the assumptions on the safety constraint s, different definitions of such an expansion operator are possible, leading to different definitions of the largest reachable safe set. For their SafeOpt algorithm, Sui et al. (2015) define such an operator leveraging the Lipschitz assumption on the safety constraint. Namely, starting from an initial set of parameters known to be safe, they recursively add to that set those parameters that must also be safe given the Lipschitz condition, until no new parameters can be added to the set. In this paper, we cannot adopt the same definition, since we do not assume Lipschitz continuity of the constraint s. We make no additional smoothness assumptions other than s being a finite norm element of the RKHS space associated with the GP kernel. Therefore, for the generalization from the current set of safe parameters we can rely solely on the GP posterior. In this setting, a natural choice for the largest safe region reachable from $\boldsymbol{x}_0$ is the set of parameters that can be classified as safe by all well behaved GPs (i.e., those GPs whose β-confidence interval contains the true value of the function) that, starting from $\boldsymbol{x}_0$, are conditioned solely on safe evaluations. This intuition leads us to the formal definition of a corresponding expansion operator R_ε and the resulting reachable safe set $S_\varepsilon(\boldsymbol{x}_0)$. This definition requires some additional objects that we now proceed to introduce. However, the intuition we provided above is sufficient to follow the rest of the paper up to the theoretical analysis, where we heavily rely on the formal definition of R_ε and $S_\varepsilon(\boldsymbol{x}_0)$. Therefore, the reader who wishes to can safely skip directly to Definition 6 and come back later to the definition of R_ε.

Formal Definition In order to formalize the intuition that the largest reachable safe set is the one that is safely reachable by all well behaved GPs, we first need to define the set of well behaved GPs that have small uncertainty over a given region of the domain. We proceed to define this set as the intersection of the set of all well behaved GPs with that of all GPs with small uncertainty over a given subset of the domain. We define these two sets in Definition 1 and Definition 2 respectively. In the following, we refer to a GP $\sim \mathcal{N}(\mu_0, k)$ conditioned on some observation data $\mathcal{D} = \{(\boldsymbol{x}_n, y_n)\}$ as $\mathrm{GP}_{\mathcal{D}}$.

Definition 1 ($\boldsymbol{\beta}$-$\mathbf{GP}_s$) *Given a kernel $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, a function $s : \mathcal{X} \to \mathbb{R}$, and $\boldsymbol{\beta} \coloneqq \{\beta_n\}$ a sequence of positive numbers, we define $\boldsymbol{\beta}$-GP_s as the set containing all the well behaved GPs, i.e., those GPs whose β-confidence interval contains the true value of s*

$$\boldsymbol{\beta}\text{-}GP_s \coloneqq \big\{ GP_{\mathcal{D}} : \ s(\boldsymbol{x}) \in [\mu_n(\boldsymbol{x}) - \beta_n \sigma_n(\boldsymbol{x}), \mu_n(\boldsymbol{x}) + \beta_n \sigma_n(\boldsymbol{x})] \forall \boldsymbol{x} \in \mathcal{X}, \forall n \leq |\mathcal{D}| \big\}. \tag{6}$$

Definition 2 ($\boldsymbol{\beta}$-$\mathbf{GP}^{\varepsilon}(\Omega)$) *Given a kernel $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, let $\Omega \subseteq \mathcal{X}$ be a subset of the kernel's domain, $\boldsymbol{\beta} \coloneqq \{\beta_n\}$ a sequence of positive numbers, and $\varepsilon > 0$. We define $\boldsymbol{\beta}$-$GP^{\varepsilon}(\Omega)$ as the set containing all the GPs that have an uncertainty over Ω smaller than ε*

$$\boldsymbol{\beta}\text{-}GP^{\varepsilon}(\Omega) \coloneqq \big\{ GP_{\mathcal{D}} : \ \beta_{|\mathcal{D}|} \sigma_{|\mathcal{D}|}(\boldsymbol{x}) \leq \varepsilon \ \forall \boldsymbol{x} \in \Omega \big\}. \tag{7}$$

Given Definitions 1 and 2, it is now straightforward to see that the intersection of $\boldsymbol{\beta}$-GP_s with $\boldsymbol{\beta}$-$\mathrm{GP}^{\varepsilon}(\Omega)$ is precisely the set of all well behaved GPs that have a small posterior uncertainty over Ω, as we formalize in Definition 3.

Definition 3 ($\boldsymbol{\beta}$-$\mathbf{GP}^{\varepsilon}_s(\Omega)$) *Given a kernel $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, a function $s : \mathcal{X} \to \mathbb{R}$, a sequence of positive numbers $\boldsymbol{\beta} \coloneqq \{\beta_n\}$, a subset of the domain $\Omega \subseteq \mathcal{X}$, and $\varepsilon > 0$, we define the set $\boldsymbol{\beta}$-$GP^{\varepsilon}_s(\Omega)$ as the set of all well behaved GPs with an uncertainty smaller than ε over Ω*

$$\boldsymbol{\beta}\text{-}GP^{\varepsilon}_s(\Omega) \coloneqq \boldsymbol{\beta}\text{-}GP_s \cap \boldsymbol{\beta}\text{-}GP^{\varepsilon}(\Omega). \tag{8}$$

If we choose the sequence $\{\beta_n\}$ as in (4), the results by Chowdhury and Gopalan (2017) imply that every $\mathrm{GP}_{\mathcal{D}}$ is well behaved with probability at least $1 - \delta$, so that β-$\mathrm{GP}^{\varepsilon}_s(\Omega)$ is not empty (see Lemma 11 in Appendix A for more details).

Now that we have a notion of well behaved GPs with localized small uncertainty, we need to define what it means for a region of the domain to be safely reachable by all such GPs. Intuitively, a set of safe parameters $\mathcal{Y} \subseteq \mathcal{X}$ is safely reachable by a GP if, starting from the safe seed $\boldsymbol{x}_0$ and progressively reducing the posterior uncertainty over the safe set, eventually $\mathcal{Y}$ is classified as safe. This intuition translates to the following expansion operator R_{ε}

Definition 4 (Expansion operator R_{ε}) *Given a Gaussian process GP, we call S_{GP} the safe set associated to such GP posterior, as prescribed by (5). With this notation, given a safe set S, $\varepsilon > 0$, and with $\boldsymbol{\beta}$-$GP^{\varepsilon}_s(S)$ as in Definition 3, we define the expansion operator R_{ε} as*

$$R_{\varepsilon}(S) \coloneqq \big\{ \boldsymbol{x} \in \mathcal{X} : \textit{for all } GP \in \boldsymbol{\beta}\text{-}GP^{\varepsilon}_s(S) \ \boldsymbol{x} \in S_{GP} \big\}, \tag{9}$$

so that $R_{\varepsilon}(S)$ is the set of all parameters that are classified as safe, according to (5), by all well behaved GPs that have an uncertainty of at most ε over S.

Algorithm 1 Generic BO loop

1: **Input:** Objective function f; GP prior s.t. $f \sim \mathrm{GP}(\mu_0, k, \sigma_\nu)$
2: **for** $n = 0, \ldots, N$ **do**
3: $\quad \boldsymbol{x}_{n+1} \leftarrow \arg\max_{\boldsymbol{x}} \alpha(\boldsymbol{x})$
4: $\quad y_{n+1} \leftarrow f(\boldsymbol{x}_{n+1}) + \nu_n$
5: $\quad$ Update GP posteriors with $(\boldsymbol{x}_{n+1}, y_{n+1})$

With the expansion operator R_ε at hand, we can finally define the largest reachable safe set as the set that we obtain by recursively applying R_ε to the safe seed $\boldsymbol{x}_0$

Definition 5 (Largest reachable safe set $S_\varepsilon(\boldsymbol{x}_0)$) *Given the expansion operator R_ε as in Definition 4, we define the largest reachable safe set starting from the safe seed $\boldsymbol{x}_0$ as the set $S_\varepsilon(\boldsymbol{x}_0)$, obtained by recursively applying R_ε to $\boldsymbol{x}_0$ an infinite number of times*

$$S_\varepsilon(\boldsymbol{x}_0) \coloneqq \lim_{n\to\infty} \underbrace{R_\varepsilon\big(R_\varepsilon\big(\ldots\big(R_\varepsilon}_{n}(\boldsymbol{x}_0)\big)\big). \tag{10}$$

The set $S_\varepsilon(\boldsymbol{x}_0)$ is, therefore, the largest set that all well behaved GPs classify as safe starting from $\boldsymbol{x}_0$, when the posterior uncertainty over the safe set at each iteration can be at most of ε.

Once we have identified the largest safe region of the domain that we can safely reach starting from the safe seed, $S_\varepsilon(\boldsymbol{x}_0)$, we can naturally define the optimum value of f within that region

Definition 6 (Safe optimum f^*_ε) *Let $\varepsilon > 0$, and $S_\varepsilon(\boldsymbol{x}_0)$ as in Definition 5. We define the safe optimum reachable from $\boldsymbol{x}_0$ as*

$$f^*_\varepsilon \coloneqq \max_{\boldsymbol{x}\in S_\varepsilon(\boldsymbol{x}_0)} f(\boldsymbol{x}). \tag{11}$$

The reachable safe optimum f^*_ε is, therefore, the natural objective for the safe optimization problem that we consider in this paper.

2.4 Bayesian Optimization

If the set $S_\varepsilon(\boldsymbol{x}_0)$ was known, then the problem of finding f^*_ε would reduce to a global optimization problem within $S_\varepsilon(\boldsymbol{x}_0)$. A successful framework to find the global optimum of an unknown and expensive function via sequential evaluations is Bayesian optimization. Given a function f defined on some fixed domain Ω, BO algorithms find the $\arg\max_{\boldsymbol{x}\in\Omega} f(\boldsymbol{x})$ by building a probabilistic model of f and evaluating parameters that maximize a so called *acquisition function* $\alpha : \Omega \to \mathbb{R}$ defined via that model. The probabilistic model is then updated in a Bayesian fashion using the evaluation results. Algorithm 1 shows pseudo-code for the generic BO loop.

There are different ways to estimate the utility associated with the evaluation at one specific parameter $\boldsymbol{x}$, which translates into a variety of acquisition functions that one can use in BO. As discussed in Section 1, however, information-theoretic acquisition functions have proven to be particularly sample efficient for both the standard (Hennig and Schuler, 2012;

Hernández-Lobato et al., 2014; Wang and Jegelka, 2017) and constrained (Perrone et al., 2019) BO problem. Based on these results, in our preliminary work (Bottero et al., 2022) we developed an information-theoretic safe exploration approach, ISE, whose derivation we recall in Section 3.1. In this paper we aim to extend the ISE exploration strategy to the full safe BO problem by combining it with a measure of the information gain about the value of the safe optimum. A well known acquisition function that exploits such a measure is the Max-value Entropy Search (MES) acquisition function (Wang and Jegelka, 2017). More specifically, the MES acquisition function computes the mutual information between an evaluation at parameter $\boldsymbol{x}$ with observed value y and the objective function's optimum value $I_n\big(\{\boldsymbol{x}, y\}; f^*\big)$, with $f^* \coloneqq \max_{\boldsymbol{x}} f(\boldsymbol{x})$, so that the next parameter to evaluate is the most informative about the value of the optimum. In the case of a GP prior and Gaussian additive noise, the MES mutual information can be expressed as

$$I_n\big(\{\boldsymbol{x}, y\}; f^*\big) = \mathbb{E}_{f^*}\left[\frac{\theta_{f^*}(\boldsymbol{x})\psi(\theta_{f^*}(\boldsymbol{x}))}{2\Psi(\theta_{f^*}(\boldsymbol{x}))} - \ln\big(\Psi(\theta_{f^*}(\boldsymbol{x}))\big)\right], \tag{12}$$

where $\theta_{f^*}(\boldsymbol{x}) = \frac{f^* - \mu_n(\boldsymbol{x})}{\sigma_n(\boldsymbol{x})}$, while ψ and Ψ are, respectively, the probability density function of a standard Gaussian and its cumulative distribution. The expectation in (12) is over the possible values of the optimum and can be approximated via a Monte Carlo estimate. Wang and Jegelka (2017) have shown that estimating (12) even with a single properly chosen sample achieves vanishing regret with high probability, meaning that, with high probability, it leads to learning about the optimum value up to arbitrary precision.

3. Safe Optimization Algorithm

In Section 2.3, we defined the goal of a safe optimization algorithm, namely finding the reachable safe optimum f^*_ε, while in Section 2.4 we introduced BO as a framework to solve unconstrained global optimization. To solve our safe optimization problem, we need to design an acquisition function $\alpha : \mathcal{X} \to \mathbb{R}$ that at each iteration n we maximize within the safe set S_n to find the next safe parameter to evaluate

$$\boldsymbol{x}_{n+1} \in \arg\max_{\boldsymbol{x} \in S_n} \alpha(\boldsymbol{x}).$$

A key element to keep in mind when designing a BO acquisition function is the exploration-exploitation trade-off, as the optimum could be both in regions where the uncertainty is low and the values seem promising as well as in areas where the uncertainty is high. In safe BO, the fact that we are only allowed to evaluate parameters that are safe with high probability adds a further layer of complexity to the exploration-exploitation problem, as applying common BO exploration strategies within the set of safe parameters is generally insufficient to drive the expansion of this set and learn about the reachable safe optimum. Sui et al. (2015) show this problematic behavior for the well known GP-UCB algorithm (Srinivas et al., 2010), while in Figure 2 we can see how the MES acquisition function (Wang and Jegelka, 2017) fails to expand the safe set sufficiently to uncover the true safe optimum when we constrain it to the current safe set.

To overcome such issue, a safe BO algorithm needs to simultaneously learn about the safety of parameters outside the current safe set, promoting its expansion, and explore-exploit

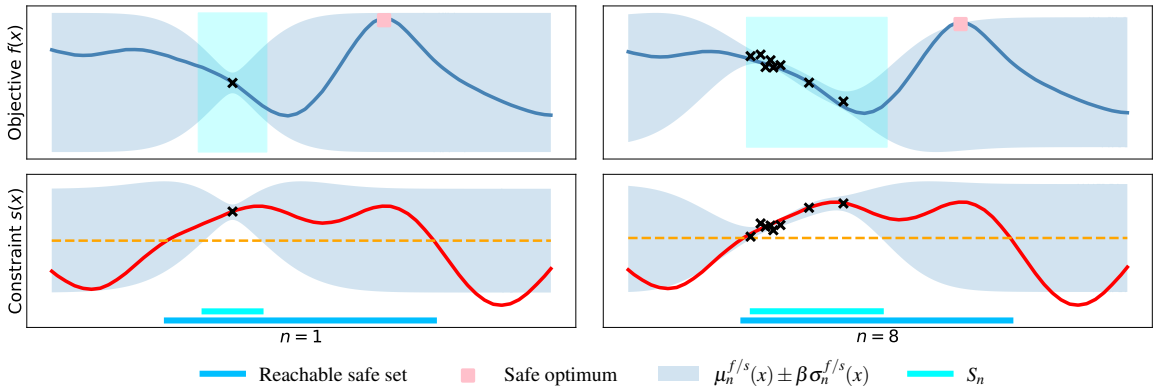


Figure 2: 1D example that illustrates why an explicit exploration component that promotes the expansion of the safe set is crucial in safe BO. Here, the function in the top half, f, is the unknown objective to optimize, while the lower half shows the unknown safety constraint, s. The algorithm that generated the plots in the figure chooses the next parameter to evaluate as the one that maximizes the pure MES acquisition function for f, constrained within the current safe set S_n. We can see that the safe optimum (pink square) is outside the initial safe set (left plots), but the acquisition function has no incentive in evaluating points that could expand the safe set further on the right side of the safe set, since the right boundary of the safe set has clearly a negligible probability of containing the optimum. Instead, the MES acquisition function focuses on the left boundary of the safe set, as we can see in the plots on the right, since that is the likely location of the optimum within the current safe set.

within the safe set to learn about the safe optimum. In order to achieve this behavior, in Section 3.1 we propose an acquisition function that selects the parameters to evaluate as the ones that maximize the information gain about the safety of parameters outside of the safe set. We then combine it with MES and show that this newly introduced acquisition function converges to the safe optimum in Section 4.

3.1 Safe Exploration

To be able to expand the safe set beyond its current extension, we need to learn what parameters outside of the safe set are also safe with high probability. Most existing safe exploration methods rely on uncertainty sampling over subsets of S_n. Instead, we exploit the ISE safe exploration strategy from (Bottero et al., 2022), that uses an information gain measure to identify the parameters that allow us to efficiently learn about the safety of parameters *outside* of S_n. That is, the ISE criterion allows us to evaluate the constraint s at safe parameters that are maximally informative about the safety of other parameters, in particular about those whose safety is uncertain. Here, we recall the derivation of the ISE acquisition function.

To start, we need a measure of information gain about the safety of parameters. We define such a measure using the binary variable $\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{x})$, which is equal to one iff $s(\boldsymbol{x}) \geq 0$

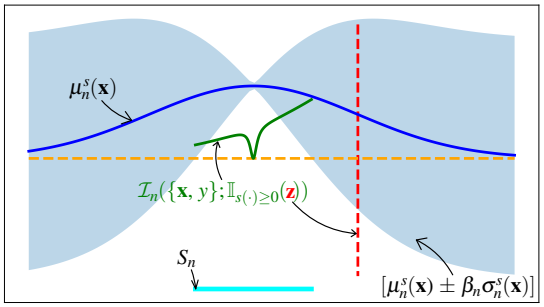


Figure 3: For a GP posterior given by the posterior mean μ_n^s and posterior variance σ_n^s, here we show in green the mutual information $I_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big)$ for different parameters $\boldsymbol{x}$ inside the safe set and for a fixed $\boldsymbol{z}$ outside (red dashed line). We see that the closer to $\boldsymbol{z}$, the more information an evaluation at $\boldsymbol{x}$ gives us about $\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})$, and that the mutual information vanishes where the posterior uncertainty is small, as there is not much left that we can learn from evaluating those parameters. The exploration part of our algorithm maximizes this quantity jointly over $\boldsymbol{x}$ and $\boldsymbol{z}$.

and zero otherwise. Its entropy according to the GP posterior at iteration n is given by

$$H_n\Big[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big] = -p_n^-(\boldsymbol{z}) \ln(p_n^-(\boldsymbol{z})) - \Big(1 - p_n^-(\boldsymbol{z})\Big) \ln(1 - p_n^-(\boldsymbol{z})), \tag{13}$$

where $p_n^-(\boldsymbol{z})$ is the probability of $\boldsymbol{z}$ being unsafe according to the GP posterior: $p_n^-(\boldsymbol{z}) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\Big(-\frac{1}{\sqrt{2}} \frac{\mu_n^s(\boldsymbol{z})}{\sigma_n^s(\boldsymbol{z})}\Big)$. The random variable $\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})$ has high-entropy where we are uncertain whether $\boldsymbol{z}$ is safe or not; that is, its entropy decreases monotonically as $|\mu_n^s(\boldsymbol{z})|$ increases and the GP posterior moves away from the safety threshold. It also decreases monotonically as $\sigma_n^s(\boldsymbol{z})$ decreases and we become more certain about whether the constraint is violated or not. This behavior also implies that the entropy goes to zero as the confidence about the safety of $\boldsymbol{z}$ increases, as desired.

Given $\mathbb{I}_{s(\cdot)\geq 0}$, we consider the mutual information $I\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big)$ between an observation y at a parameter $\boldsymbol{x}$ and the value of $\mathbb{I}_{s(\cdot)\geq 0}$ at another parameter $\boldsymbol{z}$. Since $\mathbb{I}_{s(\cdot)\geq 0}$ is the indicator function of the safe regions of the parameter space, the quantity $I_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big)$ measures how much information about the safety of $\boldsymbol{z}$ we gain by evaluating the safety constraint s at $\boldsymbol{x}$ at iteration n, averaged over all possible observed values y. This interpretation follows directly from the definition of mutual information

$$I_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big) = H_n\Big[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big] - \mathbb{E}_y\Big[H_{n+1}\Big[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big|\{\boldsymbol{x}, y\}\Big]\Big],$$

where $H_n[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})]$ is the entropy of $\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})$ according to the GP posterior at iteration n, while $H_{n+1}\Big[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big|\{\boldsymbol{x}, y\}\Big]$ is its entropy at iteration $n+1$, conditioned on a measurement y at $\boldsymbol{x}$ at iteration n. Intuitively, $I_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big)$ is negligible whenever the confidence about the safety of $\boldsymbol{z}$ is high or, more generally, whenever an evaluation at $\boldsymbol{x}$ does not have

the potential to substantially change our belief about the safety of $\boldsymbol{z}$. The mutual information is large whenever an evaluation at $\boldsymbol{x}$ on average causes the confidence about the safety of $\boldsymbol{z}$ to increase significantly. As an example, in Figure 3 we plot $I_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big)$ as a function of $\boldsymbol{x} \in S_n$ for a specific choice of $\boldsymbol{z}$ and for an RBF kernel. As one would expect, we see that the closer $\boldsymbol{x}$ gets to $\boldsymbol{z}$, the bigger the mutual information becomes, and that it vanishes in the neighborhood of previously evaluated parameters, where the posterior variance is small.

To compute $I_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big)$, we need to average (13) conditioned on an evaluation y over all possible values of y. Although we know the exact expression for the entropy of $\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})$, the integral that results from averaging over y is intractable given the expression of the entropy in (13). In order to get a tractable result, we instead derive a close approximation of (13) by truncating the Taylor expansion of $H_n[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})]$ at the second order, which appears to recover almost exactly its true behavior (see Appendix B for details)

$$H_n\Big[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big] \approx \hat{H}_n\Big[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big] \doteq \ln(2)\exp\left\{-\frac{1}{\pi\ln(2)}\left(\frac{\mu_n^s(\boldsymbol{z})}{\sigma_n^s(\boldsymbol{z})}\right)^2\right\}. \tag{14}$$

Since the posterior mean at $\boldsymbol{z}$ after an evaluation at $\boldsymbol{x}$ depends linearly on $\mu_n(\boldsymbol{x})$, and since the probability density of y depends exponentially on $-\mu_n^2(\boldsymbol{x})$, the approximation (14) reduces the conditional entropy $\mathbb{E}_y\left[\hat{H}_{n+1}\left[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\middle|\{\boldsymbol{x}, y\}\right]\right]$ to a Gaussian integral with the exact solution

$$\mathbb{E}_y\left[\hat{H}_{n+1}\left[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\middle|\{\boldsymbol{x}, y\}\right]\right] = \ln(2)\sqrt{\frac{\sigma_\nu^2 + \sigma_n^2(\boldsymbol{x})(1-\rho_n^2(\boldsymbol{x},\boldsymbol{z}))}{\sigma_\nu^2 + \sigma_n^2(\boldsymbol{x})(1+c_2\rho_n^2(\boldsymbol{x},\boldsymbol{z}))}}\exp\left\{-c_1\frac{\mu_n^2(\boldsymbol{z})}{\sigma_n^2(\boldsymbol{z})}\frac{\sigma_\nu^2 + \sigma_n^2(\boldsymbol{x})}{\sigma_\nu^2 + \sigma_n^2(\boldsymbol{x})(1+c_2\rho_n^2(\boldsymbol{x},\boldsymbol{z}))}\right\},$$

where we omitted the s superscript for μ_n, σ_n; $\rho_n(\boldsymbol{x}, \boldsymbol{z})$ is the linear correlation coefficient between $s(\boldsymbol{x})$ and $s(\boldsymbol{z})$; $c_1 \coloneqq 1/\ln(2)\pi$, and $c_2 \coloneqq 2c_1 - 1$. This result allows us to analytically calculate the approximated mutual information

$$\hat{I}_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big) \coloneqq \hat{H}_n\Big[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big] - \mathbb{E}_y\left[\hat{H}_{n+1}\left[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\middle|\{\boldsymbol{x}, y\}\right]\right]. \tag{15}$$

Note that the mutual information $I_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big)$ is symmetric, so that, alternatively, one could also choose to compute it as the average reduction in the entropy of $s(\boldsymbol{x})$ given an observation of $\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})$. However, as we show in Appendix B, the approximation (14) is very close to the true value of the entropy and computing the mutual information the other way around does not bring any additional benefit, as it also involves intractable integrals.

Now that we have defined a way to measure and compute the information gain about the safety of parameters, we can use it to design the exploration component of our acquisition function, which drives the expansion of the safe set. The natural choice is to select the parameter that maximizes the information gain; that is, we define the Information-Theoretic Safe Exploration (ISE) acquisition α^{ISE} as

$$\alpha^{\text{ISE}} \coloneqq \max_{\boldsymbol{z}\in\mathcal{X}} \hat{I}_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big). \tag{16}$$

Algorithm 2 Information-Theoretic Safe Exploration and Optimization

Input: GP priors for h (μ_0, k, σ_ν) and safe seed $\boldsymbol{x}_0$
for $n = 0, \ldots, N$ **do**
  $\boldsymbol{x}_{n+1} \leftarrow \arg\max_{\boldsymbol{x} \in S_n} \max\Big\{\alpha^{\mathrm{ISE}}(\boldsymbol{x}), \alpha^{\mathrm{MES}}(\boldsymbol{x})\Big\}$
  $y^f_{n+1}, y^s_{n+1} \leftarrow f(\boldsymbol{x}_{n+1}) + \nu^f_n,\ s(\boldsymbol{x}_{n+1}) + \nu^s_n$
  Update GP posteriors with $(\boldsymbol{x}_{n+1}, y^f_{n+1})$ and $(\boldsymbol{x}_{n+1}, y^s_{n+1})$

Evaluating s at $\boldsymbol{x} \in \arg\max_{\boldsymbol{x} \in S_n} \alpha^{\mathrm{ISE}}(\boldsymbol{x})$ maximizes the information gain about the safety of some parameter $\boldsymbol{z} \in \mathcal{X}$, so that it allows us to efficiently learn about parameters that are not yet known to be safe. While $\boldsymbol{z}$ can lie in the whole domain, the parameters whose safety we are most uncertain about lie outside the safe set. By leaving $\boldsymbol{z}$ unconstrained, we show in our theoretical analysis in Section 4 that, once we have learned about the safety of parameters outside the safe set, (16) resorts to learning about the constraint function also inside S_n. In Section 4 we also show that this behavior yields exploration guarantees, meaning that it leads to classifying the whole $S_\varepsilon(\boldsymbol{x}_0)$ as safe.

3.2 Optimization

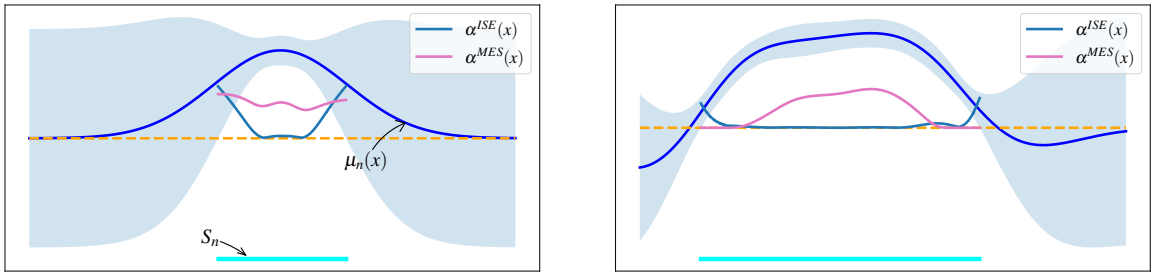
While expanding the safe set beyond its current extension is essential for the task at hand, in order to find the safe optimum f^*_ε, it is not sufficient to learn about the largest reachable safe set. Instead, one also needs to learn about the objective function values within the safe set, until the optimum can be identified with high confidence.

To this end, we can combine the mutual information $\hat{I}_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big)$ with a quantity that measures the information gain about the safe optimum value. For such combination to work, we need a quantity that has the same units as α^{ISE} to measure progress. A natural choice is therefore the mutual information between an evaluation at $\boldsymbol{x}$ and the value of the safe optimum, $I_n\big(\{\boldsymbol{x}, y\}; f^*_\varepsilon\big)$. With this choice, we can define the ISE-BO acquisition function as the maximum between the two information gain measures, so that the next parameter to evaluate, $\boldsymbol{x}_{n+1}$, is given by

$$\boldsymbol{x}_{n+1} \in \arg\max_{\boldsymbol{x} \in S_n} \alpha^{\text{ISE-BO}}(\boldsymbol{x}) \coloneqq \arg\max_{\boldsymbol{x} \in S_n} \max\Big\{\alpha^{\mathrm{ISE}}(\boldsymbol{x}), \alpha^{\mathrm{MES}}(\boldsymbol{x})\Big\}, \tag{17}$$

where $\alpha^{\mathrm{MES}} \coloneqq I_n\big(\{\boldsymbol{x}, y\}; f^*_\varepsilon\big)$ is the MES acquisition function that Wang and Jegelka (2017) proposed, as explained in Section 2.4. By selecting $\boldsymbol{x}_{n+1}$ according to (17), we choose the parameter that yields the highest information gain about the two objectives we are interested in: which parameters outside the current safe set are also safe and what is the optimum value within the current safe set. Indeed, as we show in Section 4, with high probability this selection criterion eventually leads to classifying the whole $S_\varepsilon(\boldsymbol{x}_0)$ as safe and to identifying f^*_ε with high confidence.

Figure 4 shows an example of the two components of the ISE-BO acquisition function, in a case where the safety constraint is given by $f(\boldsymbol{x}) \geq 0$. We see in Figure 4b that the exploration component α^{ISE} vanishes within the safe set, where we have high confidence about the safety of f, while it achieves its biggest values on the boundary of the safe set,



(a) $n = 2$. (b) $n = 7$.

Figure 4: Example of the two components of the ISE-BO acquisition function. For simplicity, here the safety constraint s is defined in terms of the objective function f: $s(\boldsymbol{x}) \equiv f(\boldsymbol{x})$. The two plots show the values of the safe set expansion component α^{ISE} and the optimization component α^{MES} of the acquisition function inside the safe set S_n, against the GP posterior mean and confidence interval (the blue curve and shaded area, respectively), with the safety threshold $s(\boldsymbol{x}) = f(\boldsymbol{x}) = 0$ indicated by the orange dashed line. As shown in (a), we see that at iteration $n = 2$ both components achieve their maximum on the boundary of the safe set, since, according to the GP posterior, it is both promising as a region that contains the optimum and as a region that can give us information about the safety of parameters outside S_n. The plot in (b), however, shows that, as we continue sampling on the border, α^{MES} vanishes here, since it is unlikely that the optimum is in these neighborhoods. On the contrary, α^{ISE} remains non negligible on the boundaries, as these parameters can still give us information about the safety constraint outside of S_n, possibly leading to its expansion.

where an observation is most likely to give us valuable information about the safety of parameters with uncertain safety value. On the contrary, the optimization component α^{MES} has its optimum in the inside of S_n, where the current safe optimum of f is likely to be and an evaluation would increase the amount of information about the safe optimum value. It is also possible to show that both acquisition components are bounded by a monotonically increasing function of the posterior variance, meaning that they vanish in regions where the posterior variance is very small and where, therefore, we are highly confident about the values of f and s (see Proposition 22 in Appendix A).

Figure 4 also serves as a further intuition for why the single components alone would ultimately lead to an inefficient safe optimization behavior: the exploration component keeps sampling close to the border of S_n, until the safe set cannot be expanded further, and then it tries to reduce the uncertainty on the border to very low values, before turning its attention to the inside of the safe set. On the other hand, the exploration-exploitation component within S_n drives the expansion of the safe set only until it is still plausible that the optimum is in the vicinity of its boundary. As soon as that is no longer the case, as in Figure 4b, this component keeps focusing on the inside of S_n for a large number of iterations, effectively stopping the expansion of the safe set, even if the current safe set is not yet the largest one and may not yet contain the true safe optimum f_ε^*.

Finally, in Figure 5 we show an example run of Algorithm 2. This figure shows how the acquisition function $\alpha^{\text{ISE-BO}}$ (17) alternates between evaluating parameters that lead to

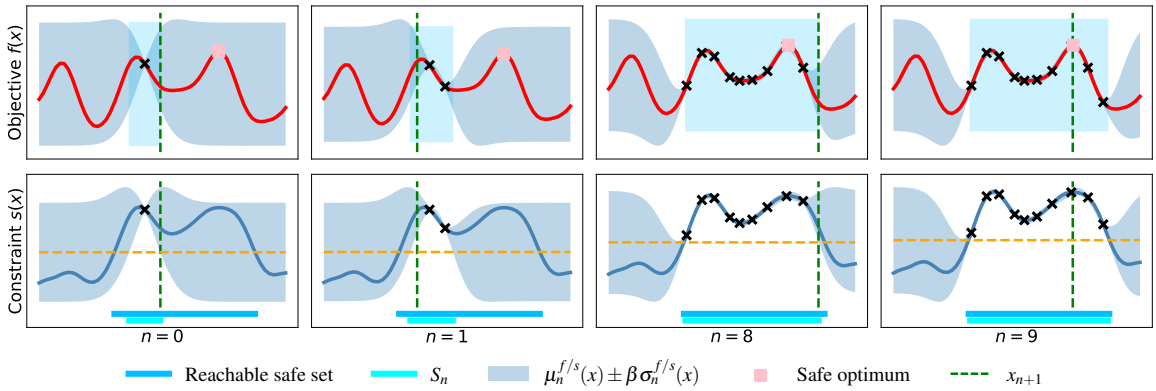


Figure 5: 1D example that illustrates Algorithm 2. The top row shows the objective function f with the corresponding posterior GP confidence interval at various iterations, while the bottom row does the same for the safety constraint s. The dashed orange line represents the safety threshold, while the black crosses mark the observed values $y_n^{f/s}$ at the evaluated parameters $\boldsymbol{x}_n$. We see how the acquisition function $\alpha^{\text{ISE-BO}}$ selects the next parameter $\boldsymbol{x}_{n+1}$ (green dashed line) alternating between parameters that are likely to expand the safe set ($n = 0$, $n = 8$) and parameters that are informative about the current safe optimum ($n = 1$, $n = 9$), until it is able to identify the reachable safe optimum (pink square).

an expansion of the safe set and parameters that give information about the safe optimum, until it confidently identifies f_ε^*. In the next section, we analyze the theoretical convergence properties of $\alpha^{\text{ISE-BO}}$ and show that it does indeed allow us to learn about the safe optimum f_ε^* with high probability.

4. Theoretical Results

In Section 2 we introduced the problem of safe BO and defined the concepts of safe set S_n and of safe optimum f_ε^*, as the largest value of the objective function f that is safely reachable from the safe seed $\boldsymbol{x}_0$. In this section, we theoretically analyze our ISE-BO acquisition function that we introduced in Section 3, and show that it leads to the discovery of the reachable safe optimum, f_ε^*. Although Algorithm 2 can be applied both to discrete and continuous domains alike, as it is common in the literature (Sui et al., 2015, 2018; Sukhija et al., 2022) in the following analysis we will restrict ourselves to a discrete and finite domain: $|\mathcal{X}| = N < \infty$. The two major results that we present here show how the acquisition function $\alpha^{\text{ISE-BO}}$ is guaranteed to find a solution of the safe BO problem. First, we show how we can find the largest reachable safe set $S_\varepsilon(\boldsymbol{x}_0)$ by sampling according to the ISE exploration component of the acquisition function. Afterwards, we show that when we combine it with MES in ISE-BO, we learn about f_ε^* to arbitrary precision. All proofs for the results presented in this section can be found in Appendix A. Before starting, we recall that all the results hold under the following assumptions

Assumption 1 *Given the objective $f : \mathcal{X} \to \mathbb{R}$ and safety constraint $s : \mathcal{X} \to \mathbb{R}$, the function $h : \mathcal{X} \times \{0, 1\} \to \mathbb{R}$ as defined in (2) has bounded norm in the RKHS space $\mathcal{H}_k$ associated*

with the GP kernel: $||h||_{\mathcal{H}_k} = B < \infty$. *Moreover, we have access to noisy observations of both functions,* $y_n^f \coloneqq f(\boldsymbol{x}_n) + \nu_n^f$ *and* $y_n^s \coloneqq s(\boldsymbol{x}_n) + \nu_n^s$, *where* $\nu_n^{f/s} \sim \mathcal{N}\Big(0, \sigma_\nu^{f/s}\Big)$.

In the following, when clear from the context, we omit the apex f/s.

Assumption 2 *Given* $\delta \in (0,1)$, *the sequence of positive numbers* $\{\beta_n\}$ *used to define the confidence intervals* $\big[\mu_n(\boldsymbol{x}, i) \pm \beta_n \sigma_n(\boldsymbol{x}, i)\big]$ *is such that the true value of* h *is contained in this interval with probability at least* $1-\delta$, *jointly for all* $(\boldsymbol{x}, i) \in \mathcal{X} \times \{0,1\}$ *and* $n \geq 1$.

As Chowdhury and Gopalan (2017) show, a choice that satisfies the condition in Assumption 2 is the one given by (4). Moreover, as mentioned in Section 1, we also note that from Assumption 2 and from the definition of the safe set (5), it follows that all parameters in S_n are safe with high probability

Proposition 7 *Choose* $\delta \in (0,1)$, *let* h *be as defined in* (2) *and the safe set* S_n *as in* (5). *Given Assumptions 1 and 2, it holds that* $s(\boldsymbol{x}) \geq 0$ *with probability of at least* $1-\delta$, *jointly for all* $\boldsymbol{x} \in S_n$ *and for all* n.

Proposition 7 guarantees that at every iteration, the parameters in the safe set are indeed safe with high probability. Now we need to show that sequentially sampling within S_n according to the $\alpha^{\text{ISE-BO}}$ acquisition function we can learn about the safe optimum.

Theorem 1 in Bottero et al. (2022) proves the preliminary result that, if the safe set stops expanding, ISE causes the uncertainty over S_n to vanish. Here we extend this result to show that by sampling according to the ISE acquisition function, we eventually learn to classify the whole largest reachable safe set $S_\varepsilon(\boldsymbol{x}_0)$ as safe.

Theorem 8 *Let* s *be the safety constraint as in Assumption 1, let* $\{\beta_n\}$ *and* δ *be as in Assumption 2, and let the domain* $\mathcal{X}$ *be discrete and of size* $D \coloneqq |\mathcal{X}| < \infty$. *Let* $\boldsymbol{x}_{n+1} \in \arg\max_{\boldsymbol{x} \in S_n} \alpha^{ISE}(\boldsymbol{x})$, *and define the true safe set* $\bar{S}(s)$ *as the set of all parameters that satisfy the safety constraint*

$$\bar{S}(s) \coloneqq \big\{\boldsymbol{x} \in \mathcal{X} : s(\boldsymbol{x}) \geq 0\big\}. \tag{18}$$

With $N_S \coloneqq |\bar{S}(s)|$, *it then holds true that for all* $\varepsilon > 0$ *there exists* N_ε *such that, with probability of at least* $1-\delta$, $S_n \supseteq S_\varepsilon(\boldsymbol{x}_0)$ *for all* $n \geq N_S N_\varepsilon$. *The smallest of such* N_ε *is given by*

$$N_\varepsilon = \min\left\{N \in \mathbb{N} : \beta_N \eta_N^{-1}\left(\frac{C\gamma_N}{N}\right) \leq \varepsilon\right\}, \tag{19}$$

where $\eta_n(x) \coloneqq \ln(2)\exp\Big\{-c_1 \frac{4\beta_n^2}{x}\Big\}\Big[1 - \sqrt{\frac{\sigma_\nu^2}{2c_1 x + \sigma_\nu^2}}\Big]$, γ_N *is the maximum information capacity of the chosen kernel, and* $C = \ln(2)/\sigma_\nu^2 \ln(1 + \sigma_\nu^{-2})$.

Theorem 8 shows that α^{ISE} is indeed a good choice for an acquisition function that promotes the expansion of the safe set, as it allows us to classify the whole reachable safe set $S_\varepsilon(\boldsymbol{x}_0)$ as safe. This result can be combined with those by Wang and Jegelka (2017), who showed that sampling according to α^{MES} in a fixed domain leads to vanishing regret with high probability. Therefore, if we first sample according to α^{ISE} until $S_{\hat{n}} \supseteq S_\varepsilon(\boldsymbol{x}_0)$, and

then, for $n > \hat{n}$, we sample according to α^{MES} in $S_{\hat{n}}$, the simple regret with respect to f^*_ε will decay to zero with high probability.

On the other hand, in general it is not necessary to learn about the whole $S_\varepsilon(\boldsymbol{x}_0)$ in order to uncover the location of the safe optimum, and it will also often be the case that the optimum is not close to the boundary of $S_\varepsilon(\boldsymbol{x}_0)$, so that it might not be necessary to learn about the whole reachable safe set in order to find it. In such situations, it could be a waste of resources to first learn about the whole $S_\varepsilon(\boldsymbol{x}_0)$ and only then turn the attention to f^*_ε. Rather, one can optimize both acquisition components at the same time, as prescribed by $\alpha^{\text{ISE-BO}}$. Theorem 10 offers convergence guarantees for this scenario, for a particular approximation of the mutual information $I_n\big(\{\boldsymbol{x}, y\}; f^*_\varepsilon\big)$.

Before we can prove Theorem 10, however, we first need the following Lemma, which presents a minor twist to a result derived by Wang and Jegelka (2017) and suggests an equivalent acquisition function to α^{MES}, when we only use one sample to compute the expectation in the mutual information $I_n\big(\{\boldsymbol{x}, y\}; f^*\big)$.

Lemma 9 *Let $\alpha^{MES}_{y^*}$ be the mutual information* (12) *where the average is computed via Monte Carlo with only one sample of $f^*_{S_n}$ of value equal to y^*. Then maximizing $\alpha^{MES}_{y^*}$ is equivalent to maximizing $\hat{\alpha}^{MES}_{y^*} \coloneqq \sigma^2_n(\boldsymbol{x})\big(y^* - \mu_n(\boldsymbol{x})\big)^{-2}$, in the sense that $\arg\max_{\boldsymbol{x}} \alpha^{MES}_{y^*}(\boldsymbol{x}) = \arg\max_{\boldsymbol{x}} \hat{\alpha}^{MES}_{y^*}(\boldsymbol{x})$.*

With the equivalent approximation $\hat{\alpha}^{\text{MES}}_{y^*}$ to the α^{MES} acquisition function, we can now show the following result

Theorem 10 *Let the domain $\mathcal{X}$ be discrete and of dimension $D \coloneqq |\mathcal{X}| < \infty$. Assume, moreover, that $\boldsymbol{x}_{n+1}$ is chosen according to $\boldsymbol{x}_{n+1} \in \arg\max \max\Big\{\alpha^{ISE}(\boldsymbol{x}), \hat{\alpha}^{MES}_{\phi}(\boldsymbol{x})\Big\}$, with $\phi \in \mathbb{R}$ such that $\phi > \mu_n(\boldsymbol{x})$ for all $\boldsymbol{x} \in S_n$ and for all n. Let, moreover, $N_S \coloneqq |\bar{S}(s)|$ be the size of the true safe set of s, and $\boldsymbol{x}^*_\varepsilon$ the reachable safe* $\arg\max$. *Then it holds that for all $\varepsilon > 0$ there exists N_ε such that, with probability of at least $1 - \delta$, $\boldsymbol{x}^*_\varepsilon \in S_n$ for all $n \geq N_S N_\varepsilon$, and $\sigma_n(\boldsymbol{x}^*_\varepsilon) \leq \varepsilon$ for all $n \geq (N_S + 1)N_\varepsilon$. The smallest of such N_ε is given by*

$$N_\varepsilon = \min\left\{ N \in \mathbb{N} : \beta_N b_N^{-1}\left(\frac{C\gamma_N}{N}\right) \leq \varepsilon \right\}, \tag{20}$$

where $b_n(x) \coloneqq \min\Big\{\eta_n(x), \frac{x}{\phi}\Big\}$, with η as given in Theorem 8, γ_N is the maximum information capacity of the chosen kernel: $\gamma_N = \max_{D\subset\mathcal{X};|D|=N} I\big(\boldsymbol{f}(D); \boldsymbol{y}(D)\big)$, and $C = \max\Big\{\frac{\ln(2)}{\sigma^2_\nu}, \frac{1}{\phi - 2\beta}\Big\}$.

Theorem 10 extends the result of Theorem 8, in that it shows that, if we evaluate parameters according to the $\alpha^{\text{ISE-BO}}$ acquisition function, then not only will we learn about $S_\varepsilon(\boldsymbol{x}_0)$, but we will also reduce the uncertainty about the optimum within it, hence solving the safe optimization task.

Both Theorems 8 and 10 find convergence bounds on n, (19) and (20), that depend on the behavior of the maximum information capacity γ_n. In Appendix A, we show that, in the discrete domain setting, (19) and (20) have a non trivial solution.

5. Discussion and Limitations

Since the ISE-BO objective is defined over a continuous domain, our proposed algorithm can deal with higher dimensional domains better than algorithms that rely on a discrete parameter space. However, because of the ISE contribution to the acquisition function, to find $\boldsymbol{x}_{n+1}$ according to $\alpha^{\text{ISE-BO}}$ we have to solve a non-convex optimization problem with twice the dimension of the parameter space, which can become computationally challenging as the dimension of the domain grows. In higher-dimensional settings, we follow LineBO by Kirschner et al. (2019), which at each iteration selects a random one-dimensional subspace to which it restricts the optimization of the acquisition function.

Our acquisition function $\alpha^{\text{ISE-BO}}$ trades off the expansion of the safe set with the search for the safe optimum. This comparison works best when both the objective function f and the constraint s share the same scale, which can be obtained without loss of generality by normalizing the two functions.

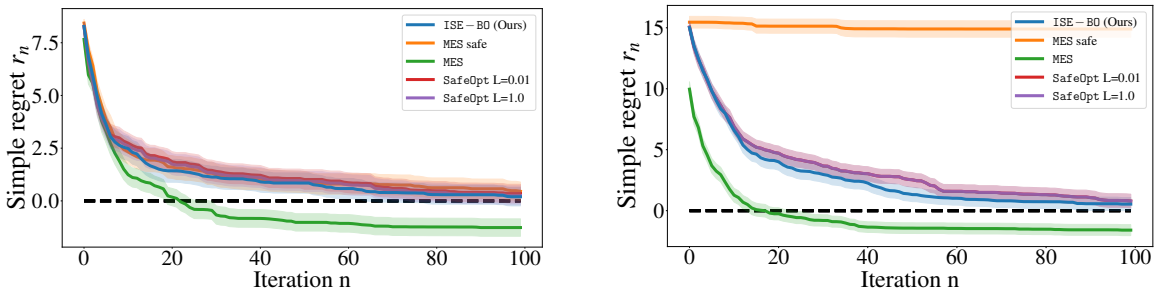
In Section 2, we assumed the observation process to be homoskedastic. However, the results can be extended to the case of heteroskedastic Gaussian noise. The observation noise at a parameter $\boldsymbol{x}$ explicitly appears in the ISE acquisition function, since it crucially affects the amount of information that we can gain by evaluating the constraint s at $\boldsymbol{x}$. On the contrary, SafeOpt-like methods do not consider the observation noise in their acquisition functions. As a consequence, ISE-BO can perform significantly better in a heteroskedastic setting, as we show in Section 6.

We also reiterate that the theoretical safety guarantees offered by ISE-BO are derived under the assumption that h is a bounded norm element of the RKHS space associated with the given GP's kernel. In applications, therefore, the choice of the kernel function becomes even more crucial when safety is an issue. For details on how to construct and choose kernels see (Garnett, 2023). The safety guarantees also depend on the choice of β_n. Typical expressions for β_n include the RKHS norm of the composite function h (Chowdhury and Gopalan, 2017; Fiedler et al., 2021), as shown in (4), which is in general difficult to estimate in practice. Moreover, to compensate for a non optimal initial choice of kernel, in practice one might also want to optimize the kernel hyperparameters during learning, based on the collected observations, with the consequence that also the RKHS norm of h would change. Because of these reasons, usually in practice a constant value of β_n is used instead.

Finally, it is worth noting that the ISE-BO acquisition function that we propose takes a myopic approach, in that it only looks at the average decrease in entropy given one observation. In standard BO, various works have shown how non-myopic approaches can achieve better performance (Yue and Kontar, 2020; Gonzalez et al., 2016; Jiang et al., 2020), and Zhang et al. (2021) argued that a non-myopic behavior is even more important in the case of constrained BO. How to extend our information based safe BO approach to the non-myopic setting is therefore an interesting direction for future work.

6. Experiments

In this section, we present experiments to evaluate the proposed ISE-BO acquisition function on different tasks and investigate its practical performance against other approaches.



(a) Same function for constraint and objective.

(b) Independent objective and constraint.

Figure 6: Performance of ISE-BO compared to SafeOpt and both constrained and unconstrained versions of MES, when the objective and constraint function are samples drawn from a GP. In both (a) and (b) we plot the average simple regret (21) and its standard error for the tested methods over 50 random seeds. In (a), the same GP sample served as both the objective f and the constraint s, while in (b), for each random seed, two different GP samples were drawn to serve, respectively, as the objective and the constraint.

GP Samples For the first part of the experiments, we evaluate ISE-BO on objective and constraint functions f and s that we obtain by sampling a GP prior at a finite number of points. This allows for an in-model comparison of ISE-BO, where we compare its performance to that of a slightly modified version of SafeOpt (Sui et al., 2015). In particular, we modify the version of SafeOpt that we use in the experiment by defining the safe set in the same way ISE-BO does, i.e., by means of the GP posterior, as done, for example, also by Berkenkamp et al. (2016b). We also compare against two versions of the MES algorithm: the standard unconstrained MES, which we allow to sample also parameters outside of the safe set, and a constrained version of it, MES-safe, that restricts the optimization of the acquisition function within the safe set. As a metric we use the simple regret, defined as

$$r_n \coloneqq \min_{\hat{n} \leq n} f^*_\varepsilon - f(\boldsymbol{x}_n), \tag{21}$$

which is the distance between the reachable safe optimum as in Definition 6 and the best parameter evaluated so far.

We select 50 independent samples of objective and constraint functions from a two-dimensional GP with an RBF kernel, defined in $[-1, 1] \times [-1, 1]$ and run the different algorithms for 100 iterations for each sample. As a first experiment, we test the setting where the safety constraint and the objective function are the same, that is $f = s$. Figure 6a shows the average simple regret over the 50 samples and the associated standard error over the regret for this setting. From this plot, we can see that ISE-BO achieves comparable results with SafeOpt and how the unconstrained version of MES finds the global optimum, which, in general, is greater than f^*_ε, leading to negative regret in the plots. We also notice that the constrained version of MES achieves a good performance, comparable with ISE-BO. This result is due to the fact that, when constraint and objective are the same randomly selected GP sample, more often than not the boundaries of the safe set happen to be promising areas

also for the location of the current safe optimum, so that also a constrained version of MES does promote expansion of the safe set, resulting in a good performance.

Subsequently, we move on to explore the setting where objective and constraint are two independently selected GP samples, and plot the corresponding results in Figure 6b. The results are very similar to the ones shown in Figure 6a, with the only noteworthy difference being the behavior of the constrained version of MES: while in the case of $f = s$ MES-safe performs on par with ISE-BO and SafeOpt, for $f \neq s$ it achieves very poor results. Such bad performance is a consequence of the safe exploration limits of pure MES when constrained to a fixed domain, as exemplified in Figure 2. In the setting $f \neq s$ these limits become relevant, since the boundary of the safe set is in general no longer an interesting location to explore for what concerns the optimum of the objective function.

Finally, in Table 1 we report the average percentage of safety violations for the various methods, and we notice that the values are comparable for all constrained methods, as the safe set is defined in the same way for all of them.

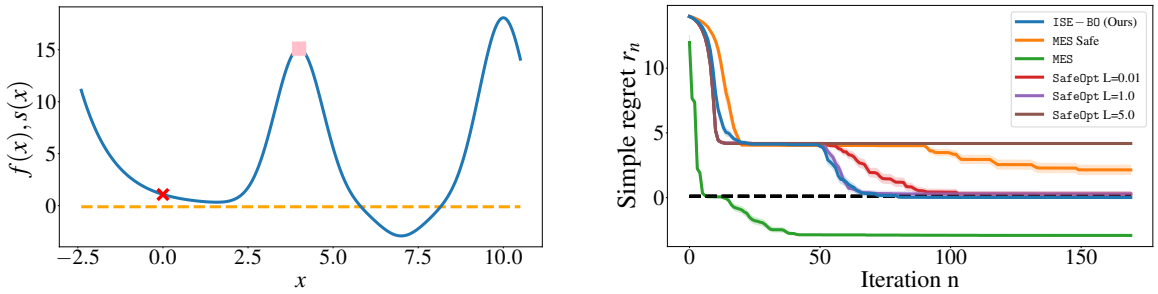
Synthetic One-Dimensional Objective To showcase the advantage of ISE-BO with respect to SafeOpt-like approaches, we tested ISE-BO, MES, MES-safe and SafeOpt for different values of the Lipschitz constant on the synthetic function shown in Figure 7a. For simplicity, this function serves both as objective and as constraint. Starting from the safe seed $\boldsymbol{x}_0 = 0$, in order to find the safe optimum the safe optimization algorithm needs to learn up to a very small error the value of the function in the region $[0, 2.3]$, where it gets very close to the safety threshold.

Once it has explored and learned about the optimum in the region $[-2.5, 0]$, ISE-BO behaves exactly this way, driven by the exploration component α^{ISE} of the acquisition function. On the other hand, SafeOpt spends time exploring that region only for appropriate values of the Lipschitz constant, as Figure 7b shows. And, naturally, MES-safe will focus on learning about the local optimum at $\boldsymbol{x} = -2.5$, away from the interesting region.

Heteroskedastic Noise Domains As discussed in Section 5, for higher dimensional domains, we can follow a similar approach to LineBO, limiting the optimization of the acquisition function to a randomly selected one-dimensional subspace of the domain. Moreover, it is also interesting to test ISE-BO in the case of heteroskedastic observation noise, since the noise is a critical quantity for our acquisition function, while it does not affect the selection criterion of SafeOpt-like methods. Therefore, in this experiment we combine a high dimensional problem with heteroskedastic noise. In particular, we apply a LineBO version of ISE-BO to the constraint and objective function $f(\boldsymbol{x}) = \frac{1}{2}e^{-||\boldsymbol{x}||^2} + e^{-||\boldsymbol{x}+\boldsymbol{x}_1||^2} + 3e^{-||\boldsymbol{x}+\boldsymbol{x}_2||^2} + e^{-||\boldsymbol{x}-\boldsymbol{x}_1||^2} + 3e^{-||\boldsymbol{x}-\boldsymbol{x}_2||^2} + 0.2$ in dimensions five and six, with the safe seed being the origin. This function has two symmetric global optima at $\boldsymbol{x}_2$ and $-\boldsymbol{x}_2$ and we set two different noise levels in the two symmetric domain halves containing the optima. To assess the exploration performance, also in this case we use the simple regret. As the results

	ISE-BO	MES safe	MES	SafeOpt L = 0.01	SafeOpt L = 1.0
% Safety violations	0.001 ± 0.002	0.004 ± 0.007	0.379 ± 0.239	0.001 ± 0.001	0.001 ± 0.001

Table 1: Average safety violations per run over the 50 runs used to obtain Figure 6



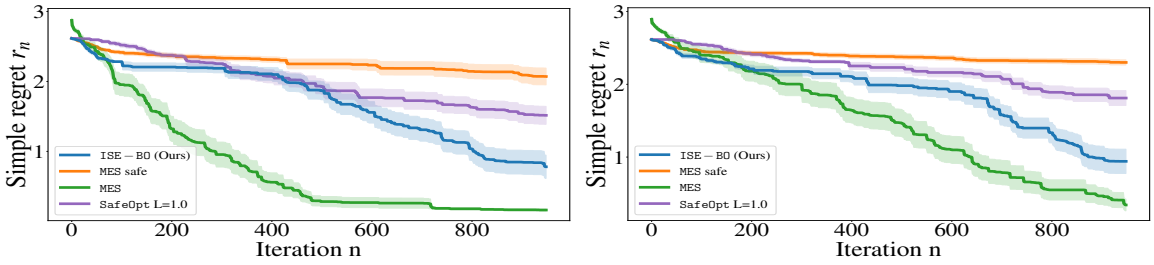
(a) One-dimensional objective-constraint.

(b) Simple regret comparison.

Figure 7: Performance of ISE-BO compared to SafeOpt and both constrained and unconstrained versions of MES for a one dimensional synthetic function, with $f = s$. In (a), we plot this function, with the value of the safe seed indicated by the red cross and the one of the safe optimum by the pink square, while in (b) we plot the average simple regret and its standard error over 50 random seeds. ISE-BO will first focus on the left of the safe seed. Once it has learned about the value of the optimum at $\boldsymbol{x} = -2.5$, it will find it more informative to explore to the right, eventually uncovering the region containing f^*_ε. On the other hand, MES-safe will spend most of its time around the local optimum at $\boldsymbol{x} = -2.5$. Similarly, SafeOpt algorithms with a too low or too high value of the Lipschitz constant will take longer to bypass the region where the constraint is very close to the safety threshold.

in Figure 8 show, ISE-BO achieves a superior sample efficiency as compared to SafeOpt, which, in turn, obtains a greater performance than the constrained version of MES, for the reasons that we discussed in Section 3. Namely, for a given number of iterations, by explicitly exploiting knowledge about the observation noise, ISE-BO is able to classify as safe regions of the domain further away from the origin, in which the objective function assumes its largest values, resulting in a smaller regret. On the other hand, SafeOpt-like methods only focus on the posterior variance, sampling parameters in both regions of the domain with similar frequency, so that the higher observation noise causes them to remain stuck in a smaller neighborhood of the origin, resulting in bigger regret.

OpenAI Gym Control After investigating the performance of ISE-BO on GP samples, we apply its exploration component to a classic control task from the OpenAI Gym framework (Brockman et al., 2016), with the goal of finding the set of parameters of a controller that satisfy some safety constraint. In particular, we consider a linear controller for the inverted pendulum. The controller is given by $u_t = x_1\theta_t + x_2\dot{\theta}_t$, where u_t is the control signal at time t, while θ_t and $\dot{\theta}_t$ are, respectively, the angular position and the angular velocity of the pendulum. Starting from a position close to the upright equilibrium, the controller's task is the stabilization of the pendulum, subject to a safety constraint on the maximum velocity reached within one episode. For a given initial controller configuration $\boldsymbol{x}_0 \coloneqq (x_1^0, x_2^0)$, we want to explore the controller's parameter space avoiding configurations that lead the pendulum to swing with a too high velocity. Namely, the safety constraint is the maximum angular velocity reached by the pendulum during an episode of fixed length, with the safety threshold set at a finite value $\dot{\theta}_M = 0.5\text{rad}/s$ that the pendulum should not exceed. In



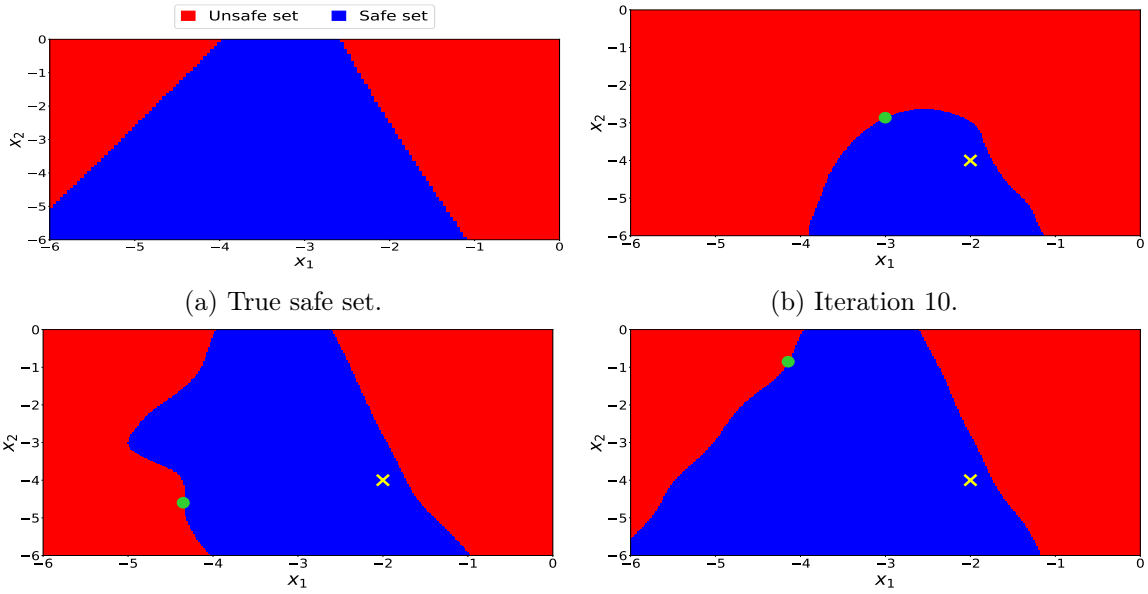
(a) Heteroskedastic noise, d = 5.

(b) Heteroskedastic noise, d = 6.

Figure 8: Performance comparison in the heteroskedastic noise experiment, for dimensions $d = 5$ and $d = 6$. We plot the simple regret r_n with respect to the safe optimum. The plots (a) and (b) show, respectively, the average regret over 30 runs in dimension five and six, as a function of the number of iterations. We can see that the LineBO version of ISE-BO promotes expansion of the safe set, leading to classifying regions where the latent function attains its largest value as safe. The plots also show that ISE-BO achieves a better sample efficiency than SafeOpt. As one would expect, we also see that the constrained version of MES achieves the worst performance, as the acquisition function has no incentive to expand the safe set. On the other hand, the unconstrained MES algorithm is naturally the quickest at identifying the optimum, since it is free to evaluate any parameter in the domain, both safe and unsafe.

Figure 9a, we show the true safe set for this problem, while in Figures 9b–9d one can see how ISE safely explores the true safe set. These plots show the parameters that ISE chooses to evaluate at different iterations. We see that in all these occasions, the ISE acquisition function focuses on the boundary of the current safe set, as this is the region that is naturally most informative about the safety of parameters outside of the safe set. This way, ISE is able to expand the safe set, until it learns the constraint function with high confidence in the vicinity of the boundary, at which point the safe set can only change slightly. This behavior eventually leads to the full true safe set being classified as safe by the GP posterior, as Figure 9d shows.

Rotary Inverted Pendulum Controller Next, we consider a simulated Furuta pendulum (Cazzolato and Prime, 2011) and an associated proportional-derivative (PD) controller, for which we want to find the optimal parameters. In particular, we consider an initial state close to the upward equilibrium state and look for parameters $\boldsymbol{x}^*$ that maintain that equilibrium with the least effort possible. The state of the pendulum is given by the four dimensional vector $\boldsymbol{\phi} \coloneqq [\theta, \alpha, \dot{\theta}, \dot{\alpha}]$, where θ is the horizontal angle, α the vertical one, and $\dot{\theta}, \dot{\alpha}$ their respective angular velocities. The objective function is given by the negative sum of the velocities and actions collected along an episode: $f(\boldsymbol{x}) \coloneqq -\sum_t \Big(|\dot{\theta}_t(\boldsymbol{x})| + |\dot{\alpha}_t(\boldsymbol{x})| + |a_t(\boldsymbol{\phi}_t; \boldsymbol{x})| \Big)$, where $\boldsymbol{x}$ are the controller parameters, while a_t is the action that the controller outputs at time-step t. To maximize this objective while maintaining the pendulum close to its upward equilibrium means to find a controller that is able to keep the pendulum up with the minimum amount of movement and effort. Closeness to the upward equilibrium position is encoded in the safety constraint: $s(\boldsymbol{x}) \coloneqq b - \sum_t |\alpha_t(\boldsymbol{x})|$, where b is the maximum allowed cumulative deviation



(a) True safe set.

(b) Iteration 10.

(c) Iteration 30.

(d) Iteration 50.

Figure 9: Safe exploration of the linear controller's parameter space in the inverted pendulum experiment. In (a) we see the true safe set, while in (b-d) we see the safe set (blue region) as learned by ISE at various iterations. The point marked by the green dot is $\boldsymbol{x}_{n+1} = \arg\max \alpha^{\mathrm{ISE}}$, while the yellow cross is the safe seed $\boldsymbol{x}_0$.

from the upward equilibrium. Positive values of s mean that α has deviated only slightly from the upward equilibrium point $\alpha = 0$ during the episode, while negative values mean a cumulative deviation larger than the allowed maximum threshold b. The PD controller has one proportional parameter and one derivative parameter associated with θ and another pair associated with α: $\boldsymbol{x} := [k_p^\theta, k_d^\theta, k_p^\alpha, k_d^\alpha]$. In this experiment, we fix k_p^θ and k_p^α, and consider the problem of finding the optimal safe k_d^θ and k_d^α. Figure 10 summarizes the results of the experiment. In this plot, we can see that also in this case ISE-BO and SafeOpt with ideal Lipschitz constant achieve a comparable result, eventually learning about the safe optimum, while the constrained version of MES struggles to efficiently explore and uncover the region of the domain that contains the safe optimum, leading to a slower convergence compared to ISE-BO. On the other hand, we can see that the unconstrained MES algorithm quickly converges to the safe optimum (which in this case is also the global one) since it is able to freely explore both safe and unsafe regions of the domain.

7. Conclusion

In this paper we considered the problem of safe Bayesian optimization. That is, given an unknown objective function f and an unknown safety constraint s, starting from an initial safe parameter $\boldsymbol{x}_0$, we want to find the safe optimum of f by only evaluating parameters that are safe with high probability, i.e., parameters $\boldsymbol{x}$ that satisfy $s(\boldsymbol{x}) \geq 0$. Since the two functions f and s are unknown and we only have access to noisy evaluations, we cannot hope

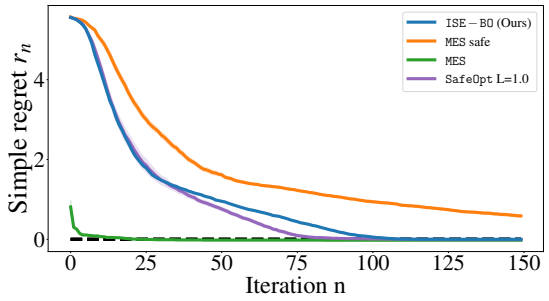


Figure 10: Performance of ISE-BO compared to SafeOpt and both constrained and unconstrained versions of MES for the Rotary inverted pendulum PD controller experiment. We plot the average over 30 random seeds and the corresponding standard error. The plot shows that ISE-BO performs comparably with SafeOpt and that both outperform the safety constrained version of MES. In this experiment, the safe optimum is also the global optimum, which explains why the green line, corresponding to the unconstrained version of MES, quickly converges to zero regret.

to find the global safe optimum if we cannot ever evaluate unsafe parameters. Instead, we assume that the safety constraint can be modeled by a GP, and use this model to define the largest set of safe parameters that are safely reachable from $\boldsymbol{x}_0$. The goal of a safe BO algorithm then becomes to identify the parameter within this set that yields the largest value of the objective f. The safety constraint exacerbates the exploration-exploitation dilemma in safe BO compared to unconstrained BO, as it is not enough to look for the optimum within the current set of safe parameters, but one also has to explicitly try to learn about the safety of parameters outside the current safe set.

To address this challenge, we have introduced ISE, an information-theoretic safe exploration criterion that maximizes information gain about the safety of parameters and promotes expansion of the safe set. In this paper, we have combined the purely exploratory ISE with Max-Value Entropy Search (Wang and Jegelka, 2017) for optimization, to obtain the novel Information-Theoretic Safe Exploration and Optimization (ISE-BO) acquisition function, $\alpha^{\text{ISE-BO}}$ (17). ISE-BO efficiently and safely explores by evaluating only parameters that are safe with high probability and by choosing those parameters that yield the greatest information gain about either the safety of other parameters or the value of the reachable safe optimum. Compared to existing approaches based on SafeOpt (Sui et al., 2015), ISE-BO does not require further hyperparameters apart from the GP model for the objective and constraint functions. Moreover, it also does not require a discretization of the domain, as it can be naturally applied to discrete and continuous domains alike.

We theoretically analyzed ISE-BO and derived two convergence results: we showed that ISE discovers the whole maximally reachable safe set, and that ISE-BO learns about the reachable safe optimum up to arbitrary precision. Our experiments support these theoretical results and demonstrate that ISE-BO performs as well as or better than SafeOpt-based methods when the best choice of Lipschitz constant is used. Future work could investigate how to further improve sample efficiency in situations where learning about the whole reachable safe set is not necessary, for example in the case where a limited budget of unsafe evaluations

is available. Another interesting direction for future work is the application of ISE-BO to real world systems, as well as the extension to the non-myopic setting.

References

Maximilian Balandat, Brian Karrer, Daniel R. Jiang, Samuel Daulton, Benjamin Letham, Andrew Gordon Wilson, and Eytan Bakshy. BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

Felix Berkenkamp, Andreas Krause, and Angela P Schoellig. Bayesian optimization with safety constraints: safe and automatic parameter tuning in robotics. *Machine Learning Journal, Special Issue on Robust Machine Learning*, 2016a.

Felix Berkenkamp, Angela P. Schoellig, and Andreas Krause. Safe controller optimization for quadrotors with gaussian processes. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2016b.

Alessandro Giacomo Bottero, Carlos E. Luis, Julia Vinogradska, Felix Berkenkamp, and Jan Peters. Information-Theoretic Safe Exploration with Gaussian Processes. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym. *arXiv preprint arXiv:1606.01540*, 2016.

Ben Cazzolato and Zebb Prime. On the dynamics of the furuta pendulum. *Hindawi Publishing Corporation Journal of Control Science and Engineering*, 2011.

Kimberly J. Chan, Joel Paulson, and Ali Mesbah. Safe explorative bayesian optimization – towards personalized treatments in plasma medicine. *IEEE Conference on Decision and Control (CDC)*, 2023.

Sayak Ray Chowdhury and Aditya Gopalan. On kernelized multi-armed bandits. In *International Conference on Machine Learning (ICML)*, 2017.

Emile Contal, Vianney Perchet, and Nicolas Vayatis. Gaussian process optimization with mutual information. In *International Conference on International Conference on Machine Learning (ICML)*, 2014.

Rikky R.P.R. Duivenvoorden, Felix Berkenkamp, Nicolas Carion, Andreas Krause, and Angela P. Schoellig. Constrained Bayesian optimization with particle swarms for safe adaptive controller tuning. In *Proc. of the IFAC (International Federation of Automatic Control)*, 2017.

Gabriel Dulac-Arnold, Daniel J. Mankowitz, and Todd Hester. Challenges of real-world reinforcement learning. In *International Conference on International Conference on Machine Learning (ICML)*, 2019.

Christian Fiedler, Carsten W. Scherer, and Sebastian Trimpe. Practical and rigorous uncertainty bounds for gaussian process regression. In *AAAI Conference on Artificial Intelligence*, 2021.

Lukas Fröhlich, Edgar Klenske, Julia Vinogradska, Christian Daniel, and Melanie Zeilinger. Noisy-input entropy search for efficient robust bayesian optimization. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020.

Roman Garnett. *Bayesian Optimization*. Cambridge University Press, 2023.

Michael A. Gelbart, Jasper Snoek, and Ryan P. Adams. Bayesian optimization with unknown constraints. In *Conference on Uncertainty in Artificial Intelligence (UAI)*, 2014.

Javier Gonzalez, Michael Osborne, and Neil Lawrence. Glasses: Relieving the myopia of bayesian optimisation. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2016.

Alkis Gotovos, Nathalie Casati, Gregory Hitz, and Andreas Krause. Active learning for level set estimation. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 2013.

Robert B. Gramacy and Herbert K. H. Lee. Optimization under unknown constraints. In *Proceedings of the Ninth Valencia International Meeting*, 2011.

Alexander Hans, Daniel Schneegass, Anton Schäfer, and Steffen Udluft. Safe exploration for reinforcement learning. In *European Symposium on Artificial Neural Networks (ESANN)*, 2008.

Philipp Hennig and Christian J. Schuler. Entropy search for information-efficient global optimization. *Journal of Machine Learning Research*, 13:1809–1837, 2012.

José Miguel Hernández-Lobato, Matthew W. Hoffman, and Zoubin Ghahramani. Predictive entropy search for efficient global optimization of black-box functions. In *International Conference on Neural Information Processing Systems (NIPS)*, 2014.

Jonas Hübotter, Bhavya Sukhija, Lenart Treven, Yarden As, and Andreas Krause. Transductive active learning: Theory and applications. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

Carl Hvarfner, Frank Hutter, and Luigi Nardi. Joint entropy search for maximally-informed bayesian optimization. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

Shali Jiang, Daniel R. Jiang, Maximilian Balandat, Brian Karrer, Jacob R. Gardner, and R. Garnett. Efficient nonmyopic bayesian optimization via one-shot multi-step trees. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 1998.

Johannes Kirschner, Mojmir Mutný, Nicole Hiller, Rasmus Ischebeck, and Andreas Krause. Adaptive and safe bayesian optimization in high dimensions via one-dimensional subspaces. In *International Conference for Machine Learning (ICML)*, 2019.

Cen-You Li, Barbara Rakitsch, and Christoph Zimmer. Safe active learning for multi-output gaussian processes. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2022.

Jonas Mockus. On bayesian methods for seeking the extremum. In *Optimization Techniques*, 1974.

Teodor Mihai Moldovan and Pieter Abbeel. Safe exploration in markov decision processes. In *International Coference on International Conference on Machine Learning (ICML)*, 2012.

Valerio Perrone, Iaroslav Shcherbatyi, Rodolphe Jenatton, Cédric Archambeau, and Matthias Seeger. Constrained bayesian optimization with max-value entropy search. In *NeurIPS 2019 Workshop on Metalearning*, 2019.

Kirill Polzounov, Ramitha Sundar, and Lee Redden. Blue river controls: A toolkit for reinforcement learning control systems on hardware. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

Manish Prajapat, Matteo Turchetta, Melanie N. Zeilinger, and Andreas Krause. Near-optimal multi-agent learning for safe coverage control. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning*. MIT Press, Cambridge, MA, USA, 2006.

Mark Schillinger, Benjamin Hartmann, Patric Skalecki, Mona Meister, Duy Nguyen-Tuong, and Oliver Nelles. Safe active learning and safe bayesian optimization for tuning a pi-controller. *IFAC-PapersOnLine*, 50(1):5967–5972, 2017.

Bernhard Schölkopf and Alexander J. Smola. *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT Press, Cambridge, MA, USA, 2002.

Matthias Schonlau, William Welch, and Donald Jones. Global versus local search in constrained optimization of computer models. *Lecture Notes–Monograph Series*, 1998.

Jens Schreiter, Duy Nguyen-Tuong, Mona Eberts, Bastian Bischoff, Heiner Markert, and Marc Toussaint. Safe exploration for active learning with gaussian processes. In *European Conference on Machine Learning and Knowledge Discovery in Databases (ECMLPKDD)*, 2015.

Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando de Freitas. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE*, 104:148–175, 2016.

Niranjan Srinivas, Andreas Krause, Sham Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *International Conference on Machine Learning (ICML)*, 2010.

Yanan Sui, Alkis Gotovos, Joel W. Burdick, and Andreas Krause. Safe exploration for optimization with Gaussian processes. In *International Conference on Machine Learning (ICML)*, 2015.

Yanan Sui, Vincent Zhuang, Joel Burdick, and Yisong Yue. Stagewise safe Bayesian optimization with Gaussian processes. In *International Conference on Machine Learning (ICML)*, 2018.

Bhavya Sukhija, Matteo Turchetta, David Lindner, Andreas Krause, Sebastian Trimpe, and Dominik Baumann. Scalable safe exploration for global optimization of dynamical systems. *CoRR*, 2022.

Matteo Turchetta, Felix Berkenkamp, and Andreas Krause. Safe exploration in finite markov decision processes with gaussian processes. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2016.

Matteo Turchetta, Felix Berkenkamp, and Andreas Krause. Safe exploration for interactive machine learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.

Yue Wang and Islam I. Hussein. *Coverage Control*, pages 11–67. Springer London, 2012.

Zi Wang and Stefanie Jegelka. Max-value entropy search for efficient bayesian optimization. In *International Conference on Machine Learning (ICML)*, 2017.

Xubo Yue and Raed Al Kontar. Why non-myopic bayesian optimization is promising and how far should we look-ahead? a study via rollout. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020.

Yunxiang Zhang, Xiangyu Zhang, and Peter I. Frazier. Two-step lookahead bayesian optimization with inequality constraints. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

Appendix A. Proofs

Before proceeding with the proof of the main results of the paper, Theorems 8 and 10, we prove Proposition 7 and Lemma 11, which show that all parameters in the safe set are indeed safe with high probability, and that the true value of the safety constraint s is contained in the posterior GP's confidence interval with high probability. In the following, when it is clear from the context, for the sake of simplicity we drop the superscripts f and s in the notation of the posterior mean and variance.

Proposition 7 *Choose $\delta \in (0,1)$, let h be as defined in (2) and the safe set S_n as in (5). Given Assumptions 1 and 2, it holds that $s(\boldsymbol{x}) \geq 0$ with probability of at least $1-\delta$, jointly for all $\boldsymbol{x} \in S_n$ and for all n.*

Proof Chowdhury and Gopalan (2017) have shown in their Theorem 2 that with the choice of $\beta_n = B + R\sqrt{2\big(\ln(e/\delta) + \gamma_n\big)}$, under the assumption of the Proposition it holds with

probability of at least $1-\delta$ that $s(\boldsymbol{x}) \in \left[\mu_n^s(\boldsymbol{x}) \pm \beta_n \sigma_n^s(\boldsymbol{x})\right]$ for all $\boldsymbol{x} \in \mathcal{X}$ and for all n, with γ_n being the maximum information gain about h after n iterations. This result, together with the way we defined the safe set in (5), implies the claim of the Proposition, concluding the proof. ■

Lemma 11 *Let s, f and h be as in (2), and let h have RKHS norm $B < \infty$. Let, moreover, $\mathcal{D} = \left\{(\boldsymbol{x}_n, y_n)\right\}_{n=1}^{|\mathcal{D}|}$ be a set of observations. Then for all $\delta \in (0,1)$, with probability of at least $1-\delta$ it holds that $s(\boldsymbol{x}) \in \left[\mu_n^s(\boldsymbol{x}) \pm \beta_n \sigma_n^s(\boldsymbol{x})\right]$ for all $\boldsymbol{x} \in \mathcal{X}$ and for all $n \leq |\mathcal{D}|$, with β_n as defined in (4).*

Proof This Lemma follows directly from the result by Chowdhury and Gopalan (2017) that we already used in the proof of Proposition 7. Specifically, Chowdhury and Gopalan (2017) have shown that with a choice of $\beta_n = B + R\sqrt{2\big(\ln(e/\delta) + \gamma_n\big)}$, it holds with probability of at least $1-\delta$ that $s(\boldsymbol{x}) \in \left[\mu_n^s(\boldsymbol{x}) \pm \beta_n \sigma_n^s(\boldsymbol{x})\right]$ for all $\boldsymbol{x} \in \mathcal{X}$ and for all n. If we now take a GP prior and a dataset $\mathcal{D}$ such that μ_n and σ_n define the posterior GP obtained by conditioning on the subset $\mathcal{D}_n = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^n \subseteq \mathcal{D}$, then the result by Chowdhury and Gopalan (2017) means that, with probability of at least $1-\delta$, $s(\boldsymbol{x}) \in \left[\mu_n^s(\boldsymbol{x}) \pm \beta_n \sigma_n^s(\boldsymbol{x})\right]$ for all $\boldsymbol{x}$ and for all $n \leq |\mathcal{D}|$. ■

Now we proceed to prove Theorems 8 and 10, that show how the ISE-BO acquisition function leads to the discovery of the whole $S_\varepsilon(\boldsymbol{x}_0)$ and of the safe optimum f_ε^*. Fundamentally, both these theorems follow from two main results. The first is that sampling according to ISE and ISE-BO on a non-expanding set $\Omega \subseteq \mathcal{X}$ causes the posterior uncertainty to vanish over the whole set Ω. The second is that if for some $\tilde{n}$ the uncertainty over the whole safe set $S_{\tilde{n}}$ is smaller than some given threshold, then S_n contains the largest reachable safe set for all $n \geq \tilde{n}$. Combining these two results with the fact that, with high probability, the safe set S_n cannot be bigger than the real total set of safe parameters $\bar{S}(s)$ yields the results of the theorems. Therefore, the roadmap to prove Theorems 8 and 10 is the following:

1. We first derive sufficient conditions for an acquisition function that imply a vanishing posterior variance when one samples according to such acquisition restricted to a non-expanding set (Lemma 12).

2. We show that, if the scaled posterior variance $\beta_n \sigma_n(\boldsymbol{x})$ is smaller than some ε over the whole safe set S_n, then the safe set includes the largest reachable safe set $S_\varepsilon(\boldsymbol{x}_0)$ (Lemma 15).

3. We derive an upper bound on the number of iterations needed for an acquisition function that satisfies the assumptions mentioned in point 1 to achieve $\beta_n \sigma_n(\boldsymbol{x}) \leq \varepsilon$ over the safe set S_n and, therefore, $S_n \supseteq S_\varepsilon(\boldsymbol{x}_0)$ (Lemma 17).

4. To prove Theorem 8, all is now left to do is to show that the ISE component of our acquisition function satisfies the conditions of Lemma 12 mentioned in point 1, which we do in Lemmas 19 and 21. Theorem 8 then follows because the safe set S_n as defined in (5) is non-shrinking and it cannot include more parameters than those present in

the real total safe set $\bar{S}(s)$, so that we can compute an upper bound on the number of iterations needed for it to stop expanding. We can then combine this upper bound with the one we mentioned in point 3 to conclude the proof of the theorem.

5. Finally, we prove Theorem 10 by showing that also the full ISE-BO acquisition satisfies the assumptions of Lemma 12, which, together with Theorem 8, imply that after at most a finite number of iterations, the uncertainty over $S_\varepsilon(\boldsymbol{x}_0)$ and, therefore, also at the reachable safe optimum will be smaller than the given threshold ε.

The remaining lemmas are technical results used to prove the lemmas mentioned in the roadmap above. We now start by proving Lemma 12, which shows that any acquisition function that satisfies some assumptions causes the posterior uncertainty to vanish, when sampling in a non-expanding set.

Lemma 12 *Let $\boldsymbol{x}_{n+1} \in \arg\max_{\boldsymbol{x}\in\mathcal{X}_n} \alpha_n(\boldsymbol{x})$, where $\mathcal{X}_n \subseteq \mathcal{X}$ is a subset of the domain of the objective and safety constraint functions, $\mathcal{X}$. If $\mathcal{X}_n$ is such that $\mathcal{X}_{n+1} \subseteq \mathcal{X}_n$ for all n, and if the acquisition function α_n satisfies the two properties:*

- *$\alpha_n(\boldsymbol{x}_{n+1}) \leq C_0\sigma_n^2(\boldsymbol{x}_{n+1})$ for some constant $C_0 > 0$;*
- *$\alpha_n(\boldsymbol{x}_{n+1}) \geq g(\tilde{\sigma}_n^2)$, with g being a monotonically increasing function, and with $\tilde{\sigma}_n^2 \coloneqq \max_{\boldsymbol{x}\in\mathcal{X}_n} \sigma_n^2(\boldsymbol{x})$,*

then it holds that for all $\varepsilon > 0$ there exists N_ε such that $\sigma_n^2(\boldsymbol{x}) \leq \varepsilon$ for every $\boldsymbol{x} \in \mathcal{X}_n$ and for all $n \geq N_\varepsilon$. The smallest of such N_ε is given by

$$N_\varepsilon = \min\left\{ N \in \mathbb{N} : g^{-1}\left(\frac{C_1\gamma_N}{N}\right) \leq \varepsilon \right\}, \tag{22}$$

where $\gamma_N = \max_{D\subset\mathcal{X};|D|=N} I\big(\boldsymbol{f}(D);\boldsymbol{y}(D)\big)$ is the maximum information capacity of the chosen kernel (Srinivas et al., 2010; Contal et al., 2014), and $C_1 = C_0/\ln(1+\sigma_\nu^{-2})$.

Proof Thanks to the assumptions on the acquisition function α_n, we know that for every n it holds that

$$\begin{aligned} g(\tilde{\sigma}_n^2) \leq \alpha_n(\boldsymbol{x}_{n+1}) &\leq C_0\sigma_n^2(\boldsymbol{x}_{n+1}) \\ \implies g(\tilde{\sigma}_n^2) &\leq C_0\sigma_n^2(\boldsymbol{x}_{n+1}) \end{aligned}$$

we can now exploit the monotonicity of g and the fact that $\tilde{\sigma}_n^2$ is not increasing if the set $\mathcal{X}_n$ does not expand with n

$$g(\tilde{\sigma}_n^2) \leq g(\tilde{\sigma}_m^2) \leq C_0\sigma_m^2(x_{m+1}) \qquad \forall n \geq m.$$

We can then use this result to write:

$$\begin{aligned} ng(\tilde{\sigma}_n^2) = \sum_{i=1}^{n} g(\tilde{\sigma}_n^2) \leq \\ C_0\sigma_\nu^2 \sum_{i=1}^{n} \sigma_\nu^{-2}\sigma_i^2(\boldsymbol{x}_{i+1}) \leq \\ \frac{C_0\sigma_\nu^2}{\sigma_\nu^2\ln(1+\sigma_\nu^{-2})} \sum_{i=1}^{n} \ln\Big(1+\sigma_\nu^{-2}\sigma_i^2(\boldsymbol{x}_{i+1})\Big) \leq C_1\gamma_n, \end{aligned} \tag{23}$$

where γ_n is the maximum information capacity and $C_1 = \frac{C_0\sigma_\nu^2}{\sigma_\nu^2 \ln(1+\sigma_\nu^{-2})}$. The second passage follows from the fact that $x \leq \ln(1+x)\sigma_\nu^{-2}/\ln(1+\sigma_\nu^{-2})$ for $x \in [0, \sigma_\nu^{-2}]$ together with the fact that $\sigma_\nu^{-2}\sigma_i^2(\boldsymbol{x}_{i+1}) \leq \sigma_\nu^{-2}k(\boldsymbol{x}_{i+1}, \boldsymbol{x}_{i+1}) \leq \sigma_\nu^{-2}$. Finally, the last passage uses the representation of the mutual information $I(\{y_n\}; \{f(\boldsymbol{x}_n)\})$ presented by Srinivas et al. (2010).

Using (23), we can show that the minimum N_ε satisfying the claim of the theorem is the one given by (22)

$$N_\varepsilon = \min\left\{ N \in \mathbb{N} : g^{-1}\left(\frac{C_1\gamma_N}{N}\right) \leq \varepsilon \right\}, \tag{24}$$

and we are now able to conclude that, as long as the information capacity grows sub-linearly in N, the set on the r.h.s. of (24) is not empty for all $\varepsilon > 0$. This is guaranteed by the fact that g^{-1} is monotonically increasing, since so is its inverse g. To check that this N_ε indeed satisfies the claim, one just has to apply g^{-1} on both initial and final state of (23) and then substitute N_ε in the place of n; the rest follows from the fact that the maximum variance is non increasing on $\mathcal{X}_n$ thanks to the assumption that $\mathcal{X}_{n+1} \subseteq \mathcal{X}_n$ for all n. ■

As anticipated above, Lemma 12 tells us that if an acquisition function satisfies the conditions of the Lemma and it is only allowed to sample within a non-expanding set $\Omega \subseteq \mathcal{X}$, then it will cause the posterior variance to vanish over Ω. Now we proceed in our roadmap to prove Theorem 8, which claims that, if we sample according to the ISE component of our acquisition function within S_n at each iteration n, then, starting from $\boldsymbol{x}_0$, with high probability we will eventually classify the whole $S_\varepsilon(\boldsymbol{x}_0)$ as safe.

Firstly, we notice that the safe set S_n as defined in (5) cannot shrink. More specifically, for every $n \geq 1$, S_n satisfies the condition $S_{n+1} \supseteq S_n$. This property means that if, for some $\tilde{n}$ the safe set $S_{\tilde{n}}$ includes the largest reachable safe set $S_\varepsilon(\boldsymbol{x}_0)$, then it will do so also for all $n > \tilde{n}$. We therefore investigate if there are any conditions that, if satisfied, guarantee that $S_\varepsilon(\boldsymbol{x}_0) \subseteq S_n$. The answer is positive, and Lemma 15 provides a sufficient condition for S_n to include $S_\varepsilon(\boldsymbol{x}_0)$. To prove Lemma 15, we first introduce a *per-GP* expansion operation:

Definition 13 (Per-GP expanded set) *Similarly to Definition 4, given a safe set S, for each $GP_\mathcal{D}$ in the set $\boldsymbol{\beta}$-$GP_s^\varepsilon(S)$, we can define the* per-GP *expanded safe set $R_\varepsilon^{GP_\mathcal{D}}$ as*

$$R_\varepsilon^{GP_\mathcal{D}} = \{\boldsymbol{x} \in \mathcal{X} : \boldsymbol{x} \in S_{GP_\mathcal{D}}\}. \tag{25}$$

Given the safe set S, the corresponding set $R_\varepsilon^{\mathrm{GP}_\mathcal{D}}$ would therefore be the set of all points classified as safe by a specific GP in $\boldsymbol{\beta}$-$\mathrm{GP}_s^\varepsilon(S)$, i.e., by a particular well behaved GP whose uncertainty is smaller than ε on S_n.

Once we recall the definition of the expansion operator R_ε (9), we see that $R_\varepsilon(S_n)$ is nothing else than the intersection of all $R_\varepsilon^{\mathrm{GP}_\mathcal{D}}$ for all $\mathrm{GP}_\mathcal{D} \in \boldsymbol{\beta}$-$\mathrm{GP}_s^\varepsilon(S_n)$, which offers a proof for the following result:

Lemma 14 *Let Ω be a subset of the domain of f and s, $\mathcal{X}$. Then for any $GP_\mathcal{D} \in \boldsymbol{\beta}$-$GP_s^\varepsilon(\Omega)$ it holds that $R_\varepsilon^{GP_\mathcal{D}} \supseteq R_\varepsilon(\Omega)$.*

Proof The claim follows directly from the definitions of the two operators $R_\varepsilon^{\mathrm{GP}_\mathcal{D}}$ and R_ε, since a parameter $\boldsymbol{x}$ that is not classified as safe by the specific $\mathrm{GP}_\mathcal{D}$ is not, by definition,

part of $R_\varepsilon(\Omega)$. ■

With the help of the set $R_\varepsilon^{\mathrm{GP}_\mathcal{D}}$ and Lemma 14, we can now easily prove Lemma 15.

Lemma 15 *Fix $\varepsilon > 0$. If $\hat{n}$ exists such that $\beta_{\hat{n}}\sigma_{\hat{n}}(\boldsymbol{x}) \leq \varepsilon \ \forall \boldsymbol{x} \in S_{\hat{n}}$, then it follows that $S_{\hat{n}} \supseteq S_\varepsilon(\boldsymbol{x}_0)$ with probability at least $1-\delta$.*

Proof We prove the claim by showing that $S_{\hat{n}} \supseteq R_\varepsilon^j(\boldsymbol{x}_0) \ \forall j$, which we can show by induction. For the base case, let's take $j = 1$ and show that $S_{\hat{n}} \supseteq R_\varepsilon(\boldsymbol{x}_0)$. This inclusion follows directly from Lemma 14, once we recall that the GP that generated $S_{\hat{n}}$, denoted as $\mathrm{GP}_{\mathcal{D}_{\hat{n}}}$, is in β-$\mathrm{GP}_s^\varepsilon(\boldsymbol{x}_0)$ with probability of at least $1-\delta$ (thanks to Lemma 11), and, therefore, $S_{\hat{n}}$ is nothing else than $R^{\mathrm{GP}_{\mathcal{D}_{\hat{n}}}}$. For the induction step, let's assume that $S_{\hat{n}} \supseteq R_\varepsilon^{j-1}(\boldsymbol{x}_0)$ and show that $S_{\hat{n}} \supseteq R_\varepsilon^j(\boldsymbol{x}_0)$. The reasoning is completely analogous to the base case with the substitution $\boldsymbol{x}_0 \to R_\varepsilon^{j-1}(\boldsymbol{x}_0)$. ■

As mentioned above, Lemma 15 gives us a condition under which the current and all future safe sets include the maximally reachable safe set $S_\varepsilon(\boldsymbol{x}_0)$. This condition is that $\beta_n\sigma_n(\boldsymbol{x}) \leq \varepsilon$ for all $\boldsymbol{x} \in S_n$. Therefore, in order to prove Theorem 8, we just need to show that there exists some n for which this condition is satisfied.

Lemma 17 below shows that such n exists for an acquisition function for which Lemma 12 holds. In order to prove Lemma 17, however, we need a couple of results about $\bar{S}(s)$, which we defined in (18) as the total true safe set $\bar{S}(s) \coloneqq \{\boldsymbol{x} \in \mathcal{X} : s(\boldsymbol{x}) \geq 0\}$. We present these results in Lemma 16.

Lemma 16 *For any $GP_\mathcal{D}$ such that $h(\boldsymbol{x}) \in [\mu_n(\boldsymbol{x}) \pm \beta_n\sigma_n(\boldsymbol{x})]$ for all $\boldsymbol{x}$ and for all $n \leq |\mathcal{D}|$, it holds that $S_n \subseteq \bar{S}(s)$. Moreover, $\forall \varepsilon > 0$ it holds that $S_\varepsilon(\boldsymbol{x}_0) \subseteq \bar{S}(s)$.*

Proof The first part of the claim is a direct consequence of the well behaved nature of the GP, as $h(\boldsymbol{x}) \in [\mu_n(\boldsymbol{x}) \pm \beta_n\sigma_n(\boldsymbol{x})]$ implies that the corresponding condition for the safety constraint s also holds: $s(\boldsymbol{x}) \in [\mu_n^s(\boldsymbol{x}) \pm \beta_n\sigma_n^s(\boldsymbol{x})]$. It therefore follows that if the lower bound of the confidence interval is above zero, so is the true value of the constraint s. The second part of the lemma follows directly from the first, since all GPs in $\boldsymbol{\beta}$-$\mathrm{GP}_s^\varepsilon\Big(R_\varepsilon^j(\boldsymbol{x}_0)\Big)$ satisfy by definition the assumption that $s(\boldsymbol{x}) \in [\mu_n^s(\boldsymbol{x}) \pm \beta_n\sigma_n^s(\boldsymbol{x})]$ for all j. ■

With Lemma 16, we can now show that by starting with $S_0 = \{\boldsymbol{x}_0\}$ and then sampling in S_n according to an acquisition function that satisfies the assumptions of Lemma 12, we will eventually classify the whole $S_\varepsilon(\boldsymbol{x}_0)$ as safe.

Lemma 17 *Let the domain $\mathcal{X}$ be discrete, and assume that $\bar{S}(s)$ contains a finite number of elements: $|\bar{S}(s)| \leq D_{\max} < \infty$. Then if we choose the parameters to evaluate as the maximizers of an acquisition function that satisfies the assumptions of Lemma 12, it holds true that $\forall \varepsilon > 0$ there exist N_ε and $n^* \leq D_{\max}N_\varepsilon$ such that, with probability of at least $1-\delta$, $S_n \supseteq S_\varepsilon(\boldsymbol{x}_0) \ \forall n \geq n^*$. The smallest N_ε is given by:*

$$N_\varepsilon = \min\left\{N \in \mathbb{N} : \beta_N g^{-1}\left(\frac{C_1\gamma_N}{N}\right) \leq \varepsilon\right\} \tag{26}$$

where γ_n, the constant C_1 and the function g are as in Lemma 12.

Proof Thanks to Lemma 12 we know that for all $\tilde{\varepsilon} > 0$, if the safe set stops expanding at iteration $\hat{n}$ it is possible to find $N_{\tilde{\varepsilon}}$ such that $\sigma_n \leq \tilde{\varepsilon}$ over the safe set $\forall n \geq \hat{n} + N_{\tilde{\varepsilon}}$. Let $N_{\varepsilon/\beta}$ be such constant for the threshold ε/β, as, for example, the one in (26), where, for simplicity, we have dropped the dependency of β on n in the notation. Because of the definition of S_n as given in (5), we have that $S_{n+1} \supseteq S_n \ \forall n$, which means that for all n either $S_{n+N_{\varepsilon/\beta}} \supset S_n$ or $S_{n+N_{\varepsilon/\beta}} \setminus S_n = \emptyset$ and, therefore, $\beta\sigma_{n+N_{\varepsilon/\beta}} \leq \varepsilon$ over the safe set. In the second case, thanks to Lemma 15 we would have that, with probability of at least $1-\delta$, $S_{n+N_{\varepsilon/\beta}} \supseteq S_\varepsilon(\boldsymbol{x}_0)$. On the other hand, in the first case, the worst case scenario is the one in which the safe set expands by a single element (i.e., $|S_{n+N_{\varepsilon/\beta}} \setminus S_n| = 1$). In that case, starting from $\boldsymbol{x}_0$, thanks to Lemma 11 and Lemma 16 we know that if we add $D_{\max}$ elements to S_n, then, with probability of at least $1-\delta$, $S_n = \bar{S}(s)$, which, thanks to Lemma 16 implies that $S_n \supseteq S_\varepsilon(\boldsymbol{x}_0)$. The proof is then complete once we recall once more that $S_{n+1} \supseteq S_n \ \forall n$. ■

With Lemma 17 at hand, in order to prove Theorem 8 we now only need to show that the exploration component of our acquisition function, α^{ISE}, satisfies the assumptions of Lemma 12. Lemma 19 shows that the ISE acquisition function satisfies the first assumption of Lemma 12, while Lemma 21 shows that it satisfies the second one. Lemmas 18 and 20 serve to prove Lemmas 19 and 21, respectively.

To simplify the notation, in the following we define $\Delta\hat{H}_n(\boldsymbol{x}, \boldsymbol{z}) = \hat{I}_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big)$.

Lemma 18 (Proved by Bottero et al. (2022) as Lemma 2) *The mutual information $\Delta\hat{H}_n(\boldsymbol{x}, \boldsymbol{z})$ as given by (15) is monotonically decreasing in $\mu_n^2(\boldsymbol{z})/\sigma_n^2(\boldsymbol{z}) \ \forall \boldsymbol{x}, \boldsymbol{z} \in \mathcal{X}$.*

Proof First of all, let us simplify notation and define $R^2 \coloneqq \mu_n^2(\boldsymbol{z})/\sigma_n^2(\boldsymbol{z})$ and $\tilde{\rho}^2 \coloneqq \rho_\nu^2(\boldsymbol{x})\rho_n^2(\boldsymbol{x}, \boldsymbol{z})$. We then need to show that

$$\frac{\partial}{\partial R^2}\left[\exp\Big\{-c_1 R^2\Big\} - \sqrt{\frac{1-\tilde{\rho}^2}{1+c_2\tilde{\rho}^2}}\exp\left\{-c_1 R^2 \frac{1}{1+c_2\tilde{\rho}^2}\right\}\right] < 0 \ \forall R, \ \forall \tilde{\rho} \in [0,1]. \tag{27}$$

We can compute the derivative and ask under which conditions it is non negative. Requiring (27) to be non negative is equivalent to ask that

$$R^2 \leq \frac{1+c_2\tilde{\rho}^2}{c_1 c_2 \tilde{\rho}^2}\left[\ln\Big(1+c_2\tilde{\rho}^2\Big) + \frac{1}{2}\ln\left(\frac{1+c_2\tilde{\rho}^2}{1-\tilde{\rho}^2}\right)\right].$$

Now, we observe that, since $c_2 \in (-1, 0)$, while $c_1 > 0$ and $\tilde{\rho}^2 \in [0, 1]$, the factor $(1 + c_2\tilde{\rho}^2)/c_1 c_2 \tilde{\rho}^2$ is always negative. For what concerns the sum of logarithms in the square brackets, it is strictly positive $\forall \tilde{\rho}^2 \in [0, 1]$, which means that, for (27) to be non negative, we would need $R^2 < 0$, which is impossible, given that $R \in \mathbb{R}$. ■

Lemma 19 (Proved by Bottero et al. (2022) as Lemma 7) *$\forall n$, $\forall \boldsymbol{x} \in S_n$, $\forall \boldsymbol{z} \in \mathcal{X}$, it holds that*

$$\Delta \hat{H}_n(\boldsymbol{x}, \boldsymbol{z}) \leq \ln(2) \frac{\sigma_n^2(\boldsymbol{x})}{\sigma_\nu^2}.$$

Proof This result can be shown directly with the following inequality chain:

$$\begin{aligned}
\hat{I}_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big) &\leq \ln(2)\left[1 - \sqrt{\frac{1-\rho_\nu^2(\boldsymbol{x})\rho_n^2(\boldsymbol{x},\boldsymbol{z})}{1+c_2\rho_\nu^2(\boldsymbol{x})\rho_n^2(\boldsymbol{x},\boldsymbol{z})}}\right] \\
&\overset{c_2\in(-1,0)}{\leq} \ln(2)\Big(1 - \sqrt{1-\rho_\nu^2(\boldsymbol{x})\rho_n^2(\boldsymbol{x},\boldsymbol{z})}\Big) \\
&\overset{\rho_\nu^2\rho_n^2\in[0,1]}{\leq} \ln(2)\Big(\rho_\nu^2(\boldsymbol{x})\rho_n^2(\boldsymbol{x},\boldsymbol{z})\Big) \\
&\overset{\rho_n^2\in[0,1]}{\leq} \ln(2)\rho_\nu^2(\boldsymbol{x}) \\
&\leq \ln(2)\frac{\sigma_n^2(\boldsymbol{x})}{\sigma_\nu^2}
\end{aligned}$$

where the first inequality follows from Lemma 18. ■

Lemma 20 (Proved by Bottero et al. (2022) as Lemma 11) *Let f be a real valued function on $\mathcal{X}$ and let μ_n and σ_n be the posterior mean and standard deviation of a GP(μ_0, k) such that there exists a non-decreasing sequence of positive numbers $\{\beta_n\}$ for which it holds that the probability $P\big\{f(\boldsymbol{x}) \in [\mu_n(\boldsymbol{x}) \pm \beta_n\sigma_n(\boldsymbol{x})]\ \forall \boldsymbol{x},\ \forall n\big\}$ is bigger than $1-\delta$. Moreover, assume that $\mu_0(\boldsymbol{x}) = 0$ for all $\boldsymbol{x}$ and that $k(\boldsymbol{x}, \boldsymbol{x}') \leq 1$ for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}$. Then it follows that $|\mu_n(\boldsymbol{x})| \leq 2\beta_n$ with probability of at least $1-\delta$ jointly for all $\boldsymbol{x}$ and for all n.*

Proof From the assumptions, it follows that the following two conditions hold for all $\boldsymbol{x}$ and for all n with probability of at least $1-\delta$:

$$|f(\boldsymbol{x})| \in \big[0, \beta_0\sigma_0(\boldsymbol{x})\big]$$

$$\mu_n(\boldsymbol{x}) \in \big[-|f(\boldsymbol{x})| - \beta_n\sigma_n(\boldsymbol{x}), |f(\boldsymbol{x})| + \beta_n\sigma_n(\boldsymbol{x})\big]$$

From these two conditions, it follows that $|\mu_n(\boldsymbol{x})| \leq \beta_0\sigma_0(\boldsymbol{x}) + \beta_n\sigma_n(\boldsymbol{x})$ with probability of at least $1-\delta$. Now, we recall that the sequence $\{\beta_n\}$ is non decreasing by assumption and that the sequence $\{\sigma_n(\boldsymbol{x})\}$ is non increasing by the properties of a GP, which allows us to conclude that $|\mu_n(\boldsymbol{x})| \leq 2\beta_n\sigma_0(\boldsymbol{x})$, which concludes the proof once we recall the assumption that $k(\boldsymbol{x}, \boldsymbol{x}') \leq 1$ for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}$. The result can easily be extended to the case of non zero prior mean, by just adding the prior mean as offset in the found upper bound for the posterior mean. ■

Lemma 21 (Based on proof by Bottero et al. (2022) of their Lemma 8) *$\forall n$, let $\tilde{\boldsymbol{x}} \in \arg\max_{S_n} \sigma_n^2(\boldsymbol{x})$ and define $\tilde{\sigma}^2 \coloneqq \sigma_n^2(\tilde{\boldsymbol{x}})$, then it holds that:*

$$\Delta\hat{H}_n(\boldsymbol{x}_{n+1}, \boldsymbol{z}_{n+1}) \geq \eta(\tilde{\sigma}^2), \tag{28}$$

with probability of at least $1-\delta$ for all $\boldsymbol{x}, \boldsymbol{z}$ and for all n, and with η given by

$$\eta(x) \coloneqq \ln(2)\exp\left\{-c_1\frac{4\beta^2}{x}\right\}\left[1-\sqrt{\frac{\sigma_\nu^2}{2c_1x+\sigma_\nu^2}}\right]. \tag{29}$$

Proof First we note that this Lemma is only non-trivial in case the posterior mean is bounded on S_n, otherwise, if we admit $|\mu_n(\boldsymbol{x})| \to \infty$, then we just recover the result that the average information gain is positive.

Now, moving to the proof, as first thing we recall that the exploration component of our algorithm always selects the $\arg\max_{\boldsymbol{x}\in S_n}$ of $\hat{I}_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big)$ as next parameter to evaluate, meaning that, by construction, $\forall \boldsymbol{x} \in S_n$ and $\forall \boldsymbol{z} \in \mathcal{X}$, it holds that:

$$\Delta\hat{H}_n(\boldsymbol{x}_{n+1}, \boldsymbol{z}_{n+1}) \geq \Delta\hat{H}_n(\boldsymbol{x}, \boldsymbol{z})$$

This implies, in particular, that $\Delta\hat{H}_n(\boldsymbol{x}_{n+1}, \boldsymbol{z}_{n+1}) \geq \Delta\hat{H}_n(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{x}})$, since, by definition, $\tilde{\boldsymbol{x}} \in S_n$ and is, therefore, always feasible. Writing this condition explicitly, we obtain:

$$\begin{aligned}
\Delta\hat{H}_n(\boldsymbol{x}_{n+1}, \boldsymbol{z}_{n+1}) &\geq \Delta\hat{H}_n(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{x}}) \\
&= \ln(2)\left[\exp\left\{-c_1\frac{\mu_n^2(\tilde{\boldsymbol{x}})}{\tilde{\sigma}^2}\right\} - \sqrt{\frac{1-\rho_\nu^2(\tilde{\boldsymbol{x}})}{1+c_2\rho_\nu^2(\tilde{\boldsymbol{x}})}}\exp\left\{-c_1\frac{\mu_n^2(\tilde{\boldsymbol{x}})}{\tilde{\sigma}^2}\frac{1}{1+c_2\rho_\nu^2(\tilde{\boldsymbol{x}})}\right\}\right] \\
&\geq \ln(2)\exp\left\{-c_1\frac{4\beta^2}{\tilde{\sigma}^2}\right\}\left[1-\sqrt{\frac{1-\rho_\nu^2(\tilde{\boldsymbol{x}})}{1+c_2\rho_\nu^2(\tilde{\boldsymbol{x}})}}\right] \\
&= \eta(\tilde{\sigma}^2)
\end{aligned}$$

where we have used the fact that $c_2 \in (-1, 0)$ and that $\rho_\nu^2(\tilde{\boldsymbol{x}}) \in [0, 1]$, in addition to Lemma 20 for the last inequality. ■

Lemmas 19 and 21 show that α^{ISE} satisfies the assumptions of Lemma 12, so that we are now able to prove Theorem 8.

Theorem 8 *Let s be the safety constraint as in Assumption 1, let $\{\beta_n\}$ and δ be as in Assumption 2, and let the domain $\mathcal{X}$ be discrete and of size $D \coloneqq |\mathcal{X}| < \infty$. Let $\boldsymbol{x}_{n+1} \in \arg\max_{\boldsymbol{x}\in S_n} \alpha^{ISE}(\boldsymbol{x})$, and define the true safe set $\bar{S}(s)$ as the set of all parameters that satisfy the safety constraint*

$$\bar{S}(s) \coloneqq \big\{\boldsymbol{x} \in \mathcal{X} : s(\boldsymbol{x}) \geq 0\big\}. \tag{18}$$

With $N_S \coloneqq |\bar{S}(s)|$, it then holds true that for all $\varepsilon > 0$ there exists N_ε such that, with probability of at least $1-\delta$, $S_n \supseteq S_\varepsilon(\boldsymbol{x}_0)$ for all $n \geq N_S N_\varepsilon$. The smallest of such N_ε is given by

$$N_\varepsilon = \min\left\{ N \in \mathbb{N} : \beta_N \eta_N^{-1}\left(\frac{C\gamma_N}{N}\right) \leq \varepsilon \right\}, \tag{19}$$

where $\eta_n(x) \coloneqq \ln(2)\exp\left\{-c_1 \frac{4\beta_n^2}{x}\right\}\left[1 - \sqrt{\frac{\sigma_\nu^2}{2c_1 x + \sigma_\nu^2}}\right]$, γ_N is the maximum information capacity of the chosen kernel, and $C = \ln(2)/\sigma_\nu^2 \ln(1+\sigma_\nu^{-2})$.

Proof From Lemmas 19 and 21 it follows that α^{ISE} satisfies the conditions of Lemma 12 with $C_0 = \ln(2)/\sigma_\nu^2$ and with $g = \eta$, as given by (29). This fact in turn implies that we can also apply Lemma 17 to it, which concludes the proof, once we substitute the correct values to C_0 and g. ■

After proving Theorem 8, we can now proceed with the proof of Theorem 10.

Theorem 10 *Let the domain $\mathcal{X}$ be discrete and of dimension $D \coloneqq |\mathcal{X}| < \infty$. Assume, moreover, that $\boldsymbol{x}_{n+1}$ is chosen according to $\boldsymbol{x}_{n+1} \in \arg\max \max\left\{\alpha^{ISE}(\boldsymbol{x}), \hat{\alpha}_\phi^{MES}(\boldsymbol{x})\right\}$, with $\phi \in \mathbb{R}$ such that $\phi > \mu_n(\boldsymbol{x})$ for all $\boldsymbol{x} \in S_n$ and for all n. Let, moreover, $N_S \coloneqq |\bar{S}(s)|$ be the size of the true safe set of s, and $\boldsymbol{x}_\varepsilon^*$ the reachable safe* $\arg\max$. *Then it holds that for all $\varepsilon > 0$ there exists N_ε such that, with probability of at least $1-\delta$, $\boldsymbol{x}_\varepsilon^* \in S_n$ for all $n \geq N_S N_\varepsilon$, and $\sigma_n(\boldsymbol{x}_\varepsilon^*) \leq \varepsilon$ for all $n \geq (N_S+1)N_\varepsilon$. The smallest of such N_ε is given by*

$$N_\varepsilon = \min\left\{ N \in \mathbb{N} : \beta_N b_N^{-1}\left(\frac{C\gamma_N}{N}\right) \leq \varepsilon \right\}, \tag{20}$$

where $b_n(x) \coloneqq \min\left\{\eta_n(x), \frac{x}{\phi}\right\}$, with η as given in Theorem 8, γ_N is the maximum information capacity of the chosen kernel: $\gamma_N = \max_{D\subset\mathcal{X};|D|=N} I\big(\boldsymbol{f}(D); \boldsymbol{y}(D)\big)$, and $C = \max\left\{\frac{\ln(2)}{\sigma_\nu^2}, \frac{1}{\phi - 2\beta}\right\}$.

Proof In order to prove the theorem, we just need to show that the acquisition function $\alpha_{\mathrm{combined}} \coloneqq \max\left\{\alpha^{\mathrm{ISE}}(\boldsymbol{x}), \hat{\alpha}_\phi^{\mathrm{MES}}(\boldsymbol{x})\right\}$ satisfies the assumptions of Lemma 12. We can then apply a reasoning analogous to the proof of Lemma 17 to show that at most after $N_S N_\varepsilon$ iterations $S_n \supseteq S_\varepsilon(\boldsymbol{x}_0)$ and, therefore, it contains $\boldsymbol{x}_\varepsilon^*$. The second part of the theorem follows by recalling the result of Lemma 16 and by noticing that the worst case scenario happens when a single new parameter is added to the safe set after exactly N_ε iterations and the last parameter to be added is precisely $\boldsymbol{x}_\varepsilon^*$. In that case after adding $\boldsymbol{x}_\varepsilon^*$ to S_n, the safe set cannot expand any more, so that after further N_ε iterations, the posterior variance will be smaller than ε over S_n and, therefore, also at $\boldsymbol{x}_\varepsilon^*$.

Using Lemma 20, it is straightforward to see that $\hat{\alpha}_\phi^{\mathrm{MES}}$ with high probability satisfies the assumptions of Lemma 12, since $\hat{\alpha}_\phi^{\mathrm{MES}} \leq \sigma_n^2(\boldsymbol{x})(\phi - 2\beta)^{-2}$, and $\hat{\alpha}_\phi^{\mathrm{MES}}(\boldsymbol{x}_{n+1}) \geq \sigma_n(\tilde{\boldsymbol{x}})/\phi$, where $\tilde{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}\in S_n} \sigma_n(\boldsymbol{x})$.

Combining these inequalities with the analogous results for α^{ISE} from Lemmas 19 and 21, we can conclude that the combined acquisition function satisfies the assumptions of Lemma 12 with

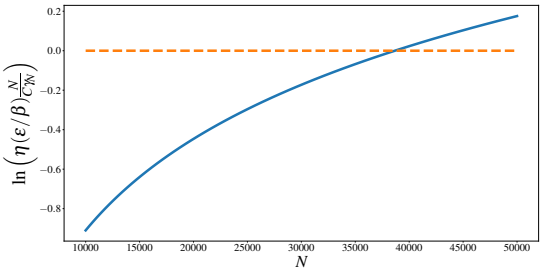


Figure 11: Asymptotic behavior of the logarithm of $\eta\big(\varepsilon/\beta_n\big)n/C\gamma_n$ (blue line), compared to $\ln(1)$ (orange dashed line). We see that for big enough n the blue line crosses the orange one, which is equivalent to the condition in (19) being satisfied.

- $C_0 = \max\Big\{\frac{\ln(2)}{\sigma_\nu^2}, \frac{1}{\phi - 2\beta}\Big\}$,
- $g(\tilde{\sigma}_n^2) = \min\Big\{\eta(\tilde{\sigma}_n^2), \frac{\tilde{\sigma}_n^2}{\phi}\Big\}$,

where η is the function defined in (29). The proof is then complete once we recall that the min of two monotonically increasing functions is also monotonically increasing. ■

As already noticed in Section 4, the convergence results of both Theorems 8 and 10 depend on the existence of N_ε as given in (19). For constant β, the existence result is trivial, given the monotonicity of the involved functions. On the other hand, when β grows with n, this is no longer obvious. However, for discrete domains and for the choice of β_n given by (4), it is also possible for N_ε to exist. Figure 11 shows an example of the asymptotic behavior of the logarithm of $\eta\big(\varepsilon/\beta_n\big)n/C\gamma_n$ for an RBF kernel. We see that, for big enough n, this logarithm exceeds the threshold of zero, marked by the orange dashed line, which is equivalent to the condition in (19) being satisfied.

What is now left to prove is Lemma 9, which shows that $\hat{\alpha}_{y^*}^{\text{MES}}$ is equivalent to $\alpha_{y^*}^{\text{MES}}$ and justifies the choice of acquisition function in Theorem 10.

Lemma 9 *Let $\alpha_{y^*}^{MES}$ be the mutual information (12) where the average is computed via Monte Carlo with only one sample of $f_{S_n}^*$ of value equal to y^*. Then maximizing $\alpha_{y^*}^{MES}$ is equivalent to maximizing $\hat{\alpha}_{y^*}^{MES} := \sigma_n^2(\boldsymbol{x})\big(y^* - \mu_n(\boldsymbol{x})\big)^{-2}$, in the sense that $\arg\max_{\boldsymbol{x}} \alpha_{y^*}^{MES}(\boldsymbol{x}) = \arg\max_{\boldsymbol{x}} \hat{\alpha}_{y^*}^{MES}(\boldsymbol{x})$.*

Proof Approximating (12) as a Monte Carlo average yields

$$I_n\big(\{\boldsymbol{x}, y\}; f^*\big) \approx \sum_{i=1}^{m}\left[\frac{\theta_{y_i^*}(\boldsymbol{x})\psi(\theta_{y_i^*}(\boldsymbol{x}))}{2\Psi(\theta_{y_i^*}(\boldsymbol{x}))} - \ln\Big(\Psi(\theta_{y_i^*}(\boldsymbol{x}))\Big)\right],$$

where $\{y_i^*\}_{i=1}^m$ are m samples of $f_{S_n}^*$. If now we restrict the sum to a single sample, we get:

$$I_n\big(\{\boldsymbol{x}, y\}; f^*\big) \approx \frac{\theta_{y^*}(\boldsymbol{x})\psi(\theta_{y^*}(\boldsymbol{x}))}{2\Psi(\theta_{y^*}(\boldsymbol{x}))} - \ln\big(\Psi(\theta_{y^*}(\boldsymbol{x}))\big). \tag{30}$$

As Wang and Jegelka (2017) show in their Lemma 3.1, maximizing (30) is equivalent to minimizing θ_{y^*}, which, in turn, is equivalent to maximizing $1/\theta_{y^*}^2$. The proof is then complete once we recall that $\theta_{y^*}(x) = \sigma_n(\boldsymbol{x})^{-1}\big(y^* - \mu_n(\boldsymbol{x})\big)$. ■

Finally, we can use Lemma 19 to also show that both mutual informations appearing in the ISE-BO acquisition function vanish for vanishing posterior variance, as mentioned in Section 3.2.

Proposition 22 *Let α^{ISE} be as defined in* (16) *and $\alpha^{MES} = I_n\big(\{\boldsymbol{x}, y\}; f_\varepsilon^*\big)$, computed using the posterior $GP_{\mathcal{D}}$ for s and for f respectively, conditioned on some data set $\mathcal{D}$, with $|\mathcal{D}| = n$. Then it holds that* $\max\Big\{\alpha^{ISE}(\boldsymbol{x}), \alpha^{MES}(\boldsymbol{x})\Big\} \leq \max\Big\{\frac{\sigma_{s,n}^2(\boldsymbol{x})}{\sigma_{s,\nu}^2}, \frac{\sigma_{f,n}^2(\boldsymbol{x})}{2\sigma_{f,\nu}^2}\Big\}$.

Proof We prove this lemma by showing that $\alpha^{\mathrm{ISE}}(\boldsymbol{x}) \leq \ln(2)\frac{\sigma_{s,n}^2(\boldsymbol{x})}{\sigma_{s,\nu}^2}$ and that $\alpha^{\mathrm{MES}}(\boldsymbol{x}) \leq \frac{1}{2}\ln\Big(1 + \sigma_{f,\nu}^{-2}\sigma_{f,n}^2(\boldsymbol{x})\Big)$. The first inequality is nothing else than the result of Lemma 19. For the second inequality, we first show that

$$I\big(f^*; \boldsymbol{y}_D\big) \leq I(f; \boldsymbol{y}_D), \tag{31}$$

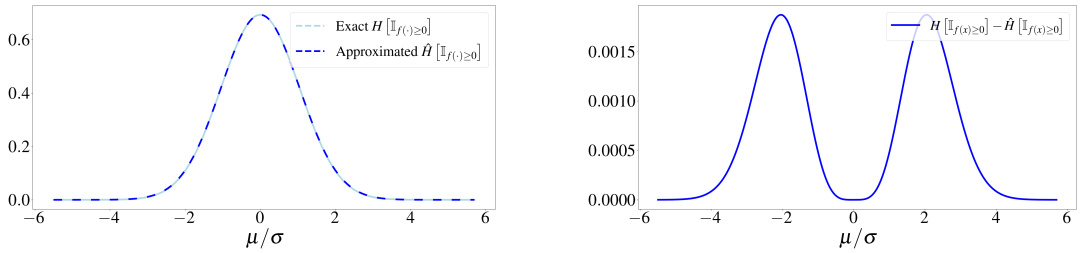
where $\boldsymbol{y}_D$ are the observed values up to iteration n: $[\boldsymbol{y}_D]_i = y_i$. To prove (31) we recall that, for any three random variables A, B, C, it holds that $I\big(A; (B, C)\big) = I(A; C) + I\big(A; B|C\big)$. From this identity, it follows that we can rewrite $I\big(\boldsymbol{y}_D; (f, f^*)\big)$ in two equivalent ways:

$$I\big(f^*; \boldsymbol{y}_D\big) + I\big(\boldsymbol{y}_D; f|f^*\big) = I\big(\boldsymbol{y}_D; (f, f^*)\big) = I(\boldsymbol{y}_D; f) + I\big(\boldsymbol{y}_D; f^*|f\big).$$

Now, if we recall the definition of f^*, it is immediate to see that $I(\boldsymbol{y}_D; f^*|f) = 0$, since f^* is completely determined by f. The result in (31) then follows from the fact that the mutual information is always non negative. If we now recall that $I_n\big(f; \{\boldsymbol{x}, y\}\big) = \frac{1}{2}\ln\Big(1 + \sigma_{f,\nu}^{-2}\sigma_{f,n}^2(\boldsymbol{x})\Big)$, we can use (31) to derive that $I_n\big(f^*; \{\boldsymbol{x}, y\}\big) \leq \frac{1}{2}\ln\Big(1 + \sigma_{f,\nu}^{-2}\sigma_{f,n}^2(\boldsymbol{x})\Big)$, which concludes the proof once we recall that $\ln(1 + x) \leq x\ \forall x > 0$ and that $\ln(2) < 1$. ■

Appendix B. Entropy of $\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})$ approximation

The mutual information $\hat{I}_n\Big(\{\boldsymbol{x}, y\}; \mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big)$ is defined in terms of $H_n\Big[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big]$, as $H_n\Big[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big] - \mathbb{E}_y\Big[H_{n+1}\Big[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big|\{\boldsymbol{x}, y\}\Big]\Big]$. In order to analytically compute it, we have approximated the entropy of the variable $\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})$ at iteration n with $\hat{H}_n\Big[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big]$, given by (14), which we have then used to derive the results presented in the paper. The approximation allowed us to derive a closed expression for the average of the entropy at parameter $\boldsymbol{z}$ after an evaluation at $\boldsymbol{x}$, $\mathbb{E}_y\Big[\hat{H}_{n+1}\Big[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\Big|\{\boldsymbol{x}, y\}\Big]\Big]$. This approximation was obtained by noticing that the exact entropy (13) has a zero mean Gaussian shape, when plotted as a function



(a) Exact and approximated $H[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})]$.

(b) Approximation error.

Figure 12: Comparison between exact entropy of $\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{x})$ (13) and approximated form (14): (a) the two entropies plotted against each other; (b) the approximation error expressed as difference of the two.

of $\mu_n(\boldsymbol{x})/\sigma_n(\boldsymbol{x})$, and then by expanding both the exact expression (13) and a zero mean un-normalized Gaussian in $\mu_n(\boldsymbol{x})/\sigma_n(\boldsymbol{x})$ in their Taylor series around zero. At the second order we obtain, respectively,

$$H_n\left[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{x})\right] = \ln(2) - \frac{1}{\pi}\left(\frac{\mu_n(\boldsymbol{x})}{\sigma_n(\boldsymbol{x})}\right)^2 + o\left(\left(\frac{\mu_n(\boldsymbol{x})}{\sigma_n(\boldsymbol{x})}\right)^2\right) \tag{32}$$

and

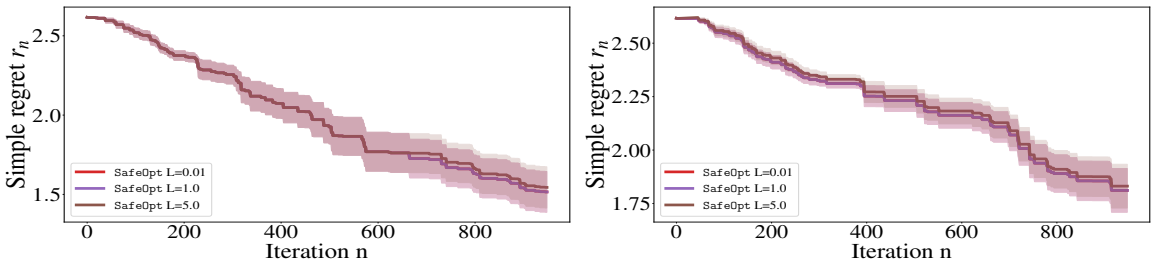
$$c_0 \exp\left\{-\frac{1}{2\sigma^2}\left(\frac{\mu_n(\boldsymbol{x})}{\sigma_n(\boldsymbol{x})}\right)^2\right\} = c_0 - c_0\frac{1}{2\sigma^2}\left(\frac{\mu_n(\boldsymbol{x})}{\sigma_n(\boldsymbol{x})}\right)^2 + o\left(\left(\frac{\mu_n(\boldsymbol{x})}{\sigma_n(\boldsymbol{x})}\right)^2\right). \tag{33}$$

By equating the terms in (32) with the ones in (33), we find $c_0 = \ln(2)$ and $\sigma^2 = \ln(2)\pi/2$, which leads to the approximation for $H_n\left[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\right]$ (14) used in the paper: $H_n\left[\mathbb{I}_{s(\cdot)\geq 0}(\boldsymbol{z})\right] \approx \ln(2)\exp\left\{-\frac{1}{\pi\ln(2)}\left(\frac{\mu_n(\boldsymbol{z})}{\sigma_n(\boldsymbol{z})}\right)^2\right\}$. In Figure 12a we plot (13) and (14) against each other as a function of the mean-standard deviation ratio, while Figure 12b shows the difference between the two. From these two plots, one can see almost perfect agreement between the two functions, with a non negligible difference limited to two small neighborhoods of the μ/σ space.

Appendix C. Experiments Details

In this Appendix, we collect details about the experiments presented in Section 6. Code for the used acquisition functions can be found at `https://github.com/boschresearch/information-theoretic-safe-exploration`. For the MES component, we used the BoTorch (Balandat et al., 2020) implementation.

GP Samples For the results shown in Figure 6 we run all methods on 50 samples from a GP defined on the square $[-1, 1]\times[-1, 1]$, with an RBF kernel with the following hyperparameters: $\mu_0 \equiv 0$; kernel lengthscale $= 0.3$; kernel outputscale $= 30$; $\sigma_\nu^2 = 0.05$, while the safe seed $\boldsymbol{x}_0$ was chosen as the origin: $\boldsymbol{x}_0 = (0, 0)$. As SafeOpt requires a discretized domain, we used the same uniform discretization of 150 points per dimension for all GP samples. Concerning



(a) Dimension 5. (b) Dimension 6.

Figure 13: Comparison of different Lipschitz constants for the SafeOpt algorithm used to generate the plots in Figure 8. We see that for this specific problem, by defining the safe set using the GP posterior, the change in Lipschitz constant does not affect significantly the performance.

the safety violations summarized in Table 1, the fact that they are comparable is expected, since in our experiments they all use the posterior GP confidence intervals to define the safe set.

Synthetic One-Dimensional Objective The objective/constraint function we used in this experiment, which we plot in Figure 7a, is given by $f(\boldsymbol{x}) = e^{-\boldsymbol{x}} + 15e^{-(\boldsymbol{x}-4)^2} + 3e^{-(\boldsymbol{x}-7)^2} + 18e^{-(\boldsymbol{x}-10)^2} + 0.41$. The GP was defined on the domain $\mathcal{X} = [-2.4, 10.5]$, with an RBF kernel with the following hyperparameters: $\mu_0 \equiv 0$; kernel lengthscale $= 0.6$; kernel outputscale $= 50$; $\sigma_\nu^2 = 0.05$, while the safe seed $\boldsymbol{x}_0$ was chosen as the origin: $\boldsymbol{x}_0 = 0$.

Heteroskedastic Noise Domains In these experiments we used the same LineBO wrapper as in the five dimensional experiment. The GP had an RBF kernel with hyperparameters: $\mu_0 \equiv 0$; kernel lengthscale $= 1.6$; kernel outputscale $= 1$; while the safe seed $\boldsymbol{x}_0$ was set to the origin. For what concerns the observation noise, as explained in Section 6, we used heteroskedastic noise, with two different values of the noise variance in the two symmetric halves of the domain. In particular, given a parameter $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$, we set the noise variance to $\sigma_\nu^2 = 0.05$ if $x_1 \geq 0$, otherwise we set it to $\sigma_\nu^2 = 0.5$. As explained in Section 6, the constraint function is $f(\boldsymbol{x}) = \frac{1}{2}e^{-\boldsymbol{x}^2} + e^{-(\boldsymbol{x}\pm\boldsymbol{x}_1)^2} + 3e^{-(\boldsymbol{x}\pm\boldsymbol{x}_2)^2} + 0.2$, with $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ given by: $\boldsymbol{x}_1 = (2.7, 0, \dots, 0)$ and $\boldsymbol{x}_2 = (6, 0, \dots, 0)$.

OpenAI Gym Control For the OpenAI Gym pendulum experiments, we used the environment provided by the OpenAI gym (Brockman et al., 2016) under the MIT license. For the safety condition, the threshold angular velocity $\dot{\theta}_M$ was set to 0.5 rad/s, with an episode length of 400 steps, and Figure 9 shows one run of ISE-BO using a GP with an RBF kernel with the following hyperparameters: $\mu_0 \equiv 0$; kernel lengthscale $= 1.3$; kernel outputscale $= 6.6$; $\sigma_\nu^2 = 0.04$.

Rotary Inverted Pendulum Controller For the Furuta pendulum experiments, we used the simulator developed by Polzounov et al. (2020) under the MIT license. For both the objective and the safety constraint, we used GP priors with an RBF kernel with the following hyperparameters: $\mu_0 \equiv 0$; kernel lengthscale $= 0.3$; kernel outputscale $= 50$; $\sigma_\nu^2 = 0.05$.