Transductive Active Learning with Application to Safe Bayesian Optimization

Jonas Hübotter*, **Bhavya Sukhija, Lenart Treven, Yarden As, Andreas Krause**
ETH Zurich, Switzerland

Abstract

Safe Bayesian optimization (Safe BO) is the task of learning an optimal policy within an unknown environment, while ensuring that safety constraints are not violated. We analyze Safe BO under the lens of a generalization of active learning with concrete prediction targets where sampling is restricted to an accessible region of the domain, while prediction targets may lie outside this region. We study a family of policies that sample adaptively to minimize uncertainty about prediction targets. We are the first to show, under general regularity assumptions, that such policies converge uniformly to the smallest possible uncertainty obtainable from the accessible data. Leveraging this result, we apply our framework to Safe BO and demonstrate that our policies improve substantially upon the state-of-the-art.

1 Introduction

Machine learning, at its core, is about designing systems that can extract knowledge or patterns from data. One part of this challenge is determining not just how to learn given observed data but deciding what data to obtain next, given the information already available. More formally, given an unknown and sufficiently regular function f over a domain $\mathcal{X}$: *How can we learn f sample-efficiently from (noisy) observations?* This problem is widely studied in *active learning* and *experimental design* (Chaloner & Verdinelli, 1995; Settles, 2009).

Active learning methods commonly aim to learn f globally, i.e., across the entire domain $\mathcal{X}$. However, in many real-world problems, **(i)** the domain is so large that learning f globally is hopeless or **(ii)** agents have limited information and cannot access the entire domain (e.g., due to restricted access or to act safely). Thus, global learning is often not desirable or even possible. Instead, intelligent systems are typically required to act in a more *directed* manner and *extrapolate* beyond their limited information. This work formalizes the above two aspects of active learning, which have remained largely unaddressed by prior work. We provide a comprehensive overview of related work in Section 5.

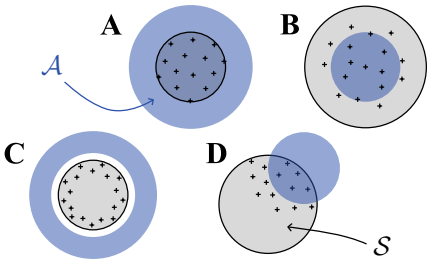

Figure 1: Instances of transductive active learning where the target space $\mathcal{A}$ is shown in blue and the sample space $\mathcal{S}$ is shown in gray. The points denote plausible observations within $\mathcal{S}$ to "learn" $\mathcal{A}$. In **(A)**, the target space contains "everything" within $\mathcal{S}$ as well as points *outside* $\mathcal{S}$. In **(B, C, D)**, one makes observations *directed* towards learning about a particular target. Prior work on active learning has focused on the instances $\mathcal{A} = \mathcal{S}$ and $\mathcal{A} \subset \mathcal{S}$.

Directed "transductive" active learning We consider the generalized problem of *transductive active learning*, where given two arbitrary subsets of the domain $\mathcal{X}$; a *target space* $\mathcal{A} \subseteq \mathcal{X}$, and a *sample space* $\mathcal{S} \subseteq \mathcal{X}$, we study the question:

How can we learn f within $\mathcal{A}$
by actively sampling observations within $\mathcal{S}$?

*Correspondence to `jonas.huebotter@inf.ethz.ch`

Workshop on Aligning Reinforcement Learning Experimentalists and Theorists (ARLET 2024).

This problem is ubiquitous in real-world applications such as safe Bayesian optimization, where $\mathcal{S}$ is a set of safe parameters and $\mathcal{A}$ might represent parameters outside $\mathcal{S}$ whose safety we want to infer. Figure 1 visualizes some instances of transductive active learning. Whereas most prior work has focused on the instance $\mathcal{X} = \mathcal{A} = \mathcal{S}$, MacKay (1992) was the first to consider specific target spaces $\mathcal{A}$ as in (B), and transductive active learning generalizes to other instances such as (A), (C), and (D).

Contributions Our main contributions are:

- We propose the principle of selecting points in $\mathcal{S}$ to minimize the "posterior uncertainty" about points in $\mathcal{A}$, which leads to natural decision rules such as ITL and VTL (Section 3).
- We are the first to give rates for the uniform convergence of uncertainty over the target space $\mathcal{A}$ to the smallest attainable value, given samples from the sample space $\mathcal{S}$ (Theorem 3.2), which implies a new generalization bound for functions in reproducing kernel Hilbert spaces (Theorem 3.3).
- We use this novel generalization bound to apply the transductive active learning framework to *safe Bayesian optimization*, an instance of safe reinforcement learning (RL). We empirically show that ITL and VTL outperform the state-of-the-art.

2 Problem Setting

We assume for now that the target space $\mathcal{A}$ and sample space $\mathcal{S}$ are finite, and relax these assumptions in the appendices. We model f as a stochastic process and denote the marginal random variables $f(\boldsymbol{x})$ by $f_{\boldsymbol{x}}$, and joint random vectors $\{f_{\boldsymbol{x}}\}_{\boldsymbol{x}\in X}$ for some $X \subseteq \mathcal{X}, |X| < \infty$ by $\boldsymbol{f}_X$. Let $\boldsymbol{y}_X$ denote the noisy observations of $\boldsymbol{f}_X$, $\{y_{\boldsymbol{x}} = f_{\boldsymbol{x}} + \varepsilon_{\boldsymbol{x}}\}_{\boldsymbol{x}\in X}$, where $\varepsilon_{\boldsymbol{x}}$ is independent noise.[2] We study the "adaptive" setting, where in round n the agent selects a point $\boldsymbol{x}_n \in \mathcal{S}$ and observes $y_n = y_{\boldsymbol{x}_n}$. The agent's choice of $\boldsymbol{x}_n$ may depend on the outcome of prior observations $\mathcal{D}_{n-1} \stackrel{\text{def}}{=} \{(\boldsymbol{x}_i, y_i)\}_{i<n}$.

Background on information theory We briefly recap several important concepts from information theory of which we provide formal definitions in Appendix B. The (differential) entropy $\mathrm{H}[\boldsymbol{f}]$ is one possible measure of uncertainty about $\boldsymbol{f}$ and the conditional entropy $\mathrm{H}[\boldsymbol{f} \mid \boldsymbol{y}]$ is the (expected) posterior uncertainty about $\boldsymbol{f}$ after observing $\boldsymbol{y}$. The information gain $\mathrm{I}(\boldsymbol{f}; \boldsymbol{y}) = \mathrm{H}[\boldsymbol{f}] - \mathrm{H}[\boldsymbol{f} \mid \boldsymbol{y}]$ measures the (expected) reduction in uncertainty about $\boldsymbol{f}$ due to $\boldsymbol{y}$. We denote the information gain about $\mathcal{A}$ from observing X by $\mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; \boldsymbol{y}_X)$. The maximum information gain about $\mathcal{A}$ from n observations within $\mathcal{S}$ is

$$\gamma_{\mathcal{A},\mathcal{S}}(n) \stackrel{\text{def}}{=} \max_{\substack{X\subseteq\mathcal{S} \\ |X|\leq n}} \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; \boldsymbol{y}_X).$$

This "information capacity" measures the information about $\boldsymbol{f}_{\mathcal{A}}$ that is accessible from within $\mathcal{S}$, and has been used previously (e.g., by Srinivas et al., 2009; Chowdhury & Gopalan, 2017; Vakili et al., 2021) in the setting where $\mathcal{X} = \mathcal{A} = \mathcal{S}$, taking the form of $\gamma_n \stackrel{\text{def}}{=} \gamma_{\mathcal{X}}(n) \stackrel{\text{def}}{=} \gamma_{\mathcal{X},\mathcal{X}}(n)$. We remark that $\gamma_{\mathcal{A},\mathcal{S}}(n) \leq \gamma_{\mathcal{S}}(n)$ holds uniformly for all $\mathcal{A}$, $\mathcal{S}$, and n due to the data processing inequality. Generally, $\gamma_{\mathcal{A},\mathcal{S}}(n)$ can be substantially smaller if the target space is a sparse subset of the sample space.

3 Main Results

We analyze the following principle for transductive active learning:

Select samples to minimize the posterior "uncertainty" about f within $\mathcal{A}$. (†)

This principle yields a family of simple and natural decision rules which depend on the chosen measure of "uncertainty". Two natural measures of uncertainty are (1) the entropy of prediction targets, $\mathrm{H}[\boldsymbol{f}_{\mathcal{A}}]$, and (2) their total variance, $\sum_{\boldsymbol{x}'\in\mathcal{A}} \mathrm{Var}[f_{\boldsymbol{x}'}]$. The corresponding decision rules are

$$(1) \qquad \boldsymbol{x}_n = \arg\min_{\boldsymbol{x}\in\mathcal{S}} \mathrm{H}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_{n-1}, y_{\boldsymbol{x}}] = \arg\max_{\boldsymbol{x}\in\mathcal{S}} \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1}), \qquad \text{(ITL)}$$

$$(2) \qquad \boldsymbol{x}_n = \arg\min_{\boldsymbol{x}\in\mathcal{S}} \operatorname{tr} \mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_{n-1}, y_{\boldsymbol{x}}] \qquad \text{(VTL)}$$

That is, ITL (short for ***I**nformation-based **T**ransductive **L**earning*) and VTL (***V**ariance-based **TL***) select $\boldsymbol{x}_n$ so as to minimize the uncertainty about the prediction targets $\boldsymbol{f}_{\mathcal{A}}$ (in expectation) after having

[2] X may be a multiset in which case repeated occurrence of $\boldsymbol{x}$ corresponds to independent observations of $y_{\boldsymbol{x}}$.

received the feedback y_n. Unlike VTL, ITL takes into account the mutual dependence between points in $\mathcal{A}$. These decision rules were suggested previously (MacKay, 1992; Seo et al., 2000; Yu et al., 2006) without deriving theoretical guarantees; and they generalize several widely used algorithms which we discuss in more detail in Section 5. Most prominently, in the global learning setting where $\mathcal{S} \subseteq \mathcal{A}$, ITL reduces to $\boldsymbol{x}_n = \arg\max_{\boldsymbol{x} \in \mathcal{S}} \mathrm{I}(f_{\boldsymbol{x}}; y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1})$, i.e., is "undirected" and reduces to standard uncertainty-based active learning strategies (cf. Appendix C.1). The convergence properties for the special instance of ITL with $\mathcal{S} = \mathcal{A}$ have been studied extensively. To the best of our knowledge, we are the first to extend these guarantees to the more general setting of transductive active learning.

In our presented results, we make the following assumption.

Assumption 3.1. In the case of ITL, the information gain $\psi_{\mathcal{A}}(X) = \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; \boldsymbol{y}_X)$ is submodular. In the case of VTL, the variance reduction $\psi_{\mathcal{A}}(X) = \mathrm{tr}\ \mathrm{Var}[\boldsymbol{f}_{\mathcal{A}}] - \mathrm{tr}\ \mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_X]$ is submodular.

Under this assumption, $\psi_{\mathcal{A}}(\boldsymbol{x}_{1:n})$ is a constant factor approximation of $\max_{X \subseteq \mathcal{S}, |X| \leq n} \psi_{\mathcal{A}}(X)$ due to the seminal result on submodular function maximization by Nemhauser et al. (1978). Similar assumptions have been made, e.g., by Bogunovic et al. (2016) and Kothawade et al. (2021). Assumption 3.1 is satisfied exactly for ITL when $\mathcal{S} \subseteq \mathcal{A}$ and f is a Gaussian process (cf. Lemma C.9), and we provide an extensive discussion of our results in Appendix C.3 for instances where Assumption 3.1 is satisfied approximately, relying on the notion of weak submodularity (Das & Kempe, 2018).

3.1 Gaussian Process Setting

When $f \sim \mathcal{GP}(\mu, k)$ is a Gaussian process (GP, Williams & Rasmussen, 2006) with known mean function μ and kernel k, and the noise $\varepsilon_{\boldsymbol{x}}$ is mutually independent and zero-mean Gaussian with known variance, the ITL and VTL objectives have a closed form expression (cf. Appendix F). Further, the information capacity γ_n is sublinear in n for a rich class of GPs (Srinivas et al., 2009; Vakili et al., 2021), with rates summarized in Table 1 of the appendix.

Convergence to irreducible uncertainty So far, our discussion was centered around the role of the target space $\mathcal{A}$ in facilitating *directed* learning. An orthogonal contribution of this work is to study *extrapolation* from the sample space $\mathcal{S}$ to points $\boldsymbol{x} \in \mathcal{A} \setminus \mathcal{S}$. To this end, we derive bounds on the marginal posterior variance $\sigma_n^2(\boldsymbol{x}) \stackrel{\text{def}}{=} \mathrm{Var}[f(\boldsymbol{x}) \mid \mathcal{D}_n]$ for points in $\mathcal{A}$. These bounds depend on the instance of transductive active learning (i.e., $\mathcal{A}$ and $\mathcal{S}$) and might be of independent interest for active learning. For ITL and VTL, they imply uniform convergence of the variance for a rich class of GPs. To the best of our knowledge, this work is the first to present such bounds.

We define the *irreducible uncertainty* as the variance of $f(\boldsymbol{x})$ provided complete knowledge of f in $\mathcal{S}$:

$$\eta_{\mathcal{S}}^2(\boldsymbol{x}) \stackrel{\text{def}}{=} \mathrm{Var}[f_{\boldsymbol{x}} \mid \boldsymbol{f}_{\mathcal{S}}].$$

As the name suggests, $\eta_{\mathcal{S}}^2(\boldsymbol{x})$ represents the smallest uncertainty one can hope to achieve from observing only within $\mathcal{S}$. For all $\boldsymbol{x} \in \mathcal{S}$, it is easy to see that $\eta_{\mathcal{S}}^2(\boldsymbol{x}) = 0$. However, the irreducible uncertainty of $\boldsymbol{x} \notin \mathcal{S}$ may be (and typically is!) strictly positive.

Theorem 3.2 (Bound on marginal variance for ITL and VTL)**.** *Let Assumption 3.1 hold and the data be selected by either* ITL *or* VTL*.. Assume that $f \sim \mathcal{GP}(\mu, k)$ with known mean function μ and kernel k, the noise $\varepsilon_{\boldsymbol{x}}$ is mutually independent and zero-mean Gaussian with known variance, and γ_n is sublinear in n. Then there exists a constant C such that for any $n \geq 1$ and $\boldsymbol{x} \in \mathcal{A}$,*

$$\sigma_n^2(\boldsymbol{x}) \leq \underbrace{\eta_{\mathcal{S}}^2(\boldsymbol{x})}_{\textit{irreducible}} + \underbrace{C \frac{\gamma_{\mathcal{A},\mathcal{S}}(n)}{\sqrt{n}}}_{\textit{reducible}}. \tag{1}$$

Moreover, if $\boldsymbol{x} \in \mathcal{A} \cap \mathcal{S}$, there exists a constant C' such that

$$\sigma_n^2(\boldsymbol{x}) \leq C' \frac{\gamma_{\mathcal{A},\mathcal{S}}(n)}{n}. \tag{2}$$

Intuitively, Equation (1) of Theorem 3.2 can be understood as bounding an epistemic "generalization gap" (Wainwright, 2019) of the learner. The reducible uncertainty converges to zero, e.g., for Gaussian and smooth Matérn kernels, at all prediction targets $\boldsymbol{x} \in \mathcal{A}$. As to be expected, a smaller target space (i.e., more targeted sampling) leads to faster convergence due to a smaller information capacity $\gamma_{\mathcal{A},\mathcal{S}}(n)$. Equation (2) matches prior results for the setting $\mathcal{S} = \mathcal{A}$. We provide a formal proof of Theorem 3.2 in Appendix C.5.

3.2 Agnostic Setting

The result from the GP setting translates also to the agnostic setting, where the "ground truth" $f^\star$ may be any sufficiently regular fixed function on $\mathcal{X}$.[3] In this case, we use the model f from Section 3.1 as a (misspecified) model of $f^\star$, with some kernel k and zero mean function $\mu(\cdot) = 0$. We denote by $\mu_n(\boldsymbol{x}) = \mathbb{E}[f(\boldsymbol{x}) \mid \mathcal{D}_n]$ the posterior mean of f. W.l.o.g. we assume in the following result that the prior variance is bounded, i.e., $\mathrm{Var}[f(\boldsymbol{x})] \leq 1$.

Theorem 3.3 (Bound on approximation error for ITL and VTL, following Abbasi-Yadkori (2013); Chowdhury & Gopalan (2017))**.** *Let Assumption 3.1 hold and the data be selected by either* ITL *or* VTL*. Pick any $\delta \in (0,1)$. Assume that $f^\star$ lies in the reproducing kernel Hilbert space $\mathcal{H}_k(\mathcal{X})$ of the kernel k with norm $\|f^\star\|_k < \infty$, the noise ε_n is conditionally ρ-sub-Gaussian, and γ_n is sublinear in n. Let $\beta_n(\delta) = \|f^\star\|_k + \rho\sqrt{2(\gamma_n + 1 + \log(1/\delta))}$. Then for any $n \geq 1$ and $\boldsymbol{x} \in \mathcal{A}$, jointly with probability at least $1-\delta$,*

$$|f^\star(\boldsymbol{x}) - \mu_n(\boldsymbol{x})| \leq \beta_n(\delta)\Big[\underbrace{\eta_{\mathcal{S}}(\boldsymbol{x})}_{\textit{irreducible}} + \underbrace{\nu_{\mathcal{A},\mathcal{S}}(n)}_{\textit{reducible}}\Big]$$

where $\nu^2_{\mathcal{A},\mathcal{S}}(n)$ denotes the reducible part of Equation (1).

We provide a formal proof of Theorem 3.3 in Appendix C.6. Theorem 3.3 generalizes approximation error bounds of prior works to the extrapolation setting, where some prediction targets $\boldsymbol{x} \in \mathcal{A}$ lie outside the sample space $\mathcal{S}$. For prediction targets $\boldsymbol{x} \in \mathcal{A} \cap \mathcal{S}$, the irreducible uncertainty vanishes, and we recover previous results from the setting $\mathcal{S} = \mathcal{A}$.

Theorems 3.2 and 3.3 show that ITL and VTL efficiently learn f at the prediction targets $\mathcal{A}$ for large classes of "sufficiently regular" functions f. In the following, we validate these results experimentally by showing that ITL and VTL exhibit strong empirical performance in a broad range of applications.

3.3 Experiments in the Gaussian Process Setting

How does the smoothness of f affect ITL? We contrast two "extreme" kernels: the *Gaussian kernel* $k(\boldsymbol{x}, \boldsymbol{x'}) = \exp(-\|\boldsymbol{x} - \boldsymbol{x'}\|_2^2/2)$ and the *Laplace kernel* $k(\boldsymbol{x}, \boldsymbol{x'}) = \exp(-\|\boldsymbol{x} - \boldsymbol{x'}\|_1)$. In the mean-squared sense, the Gaussian kernel yields a smooth process f whereas the Laplace kernel yields a continuous but non-differentiable f (Williams & Rasmussen, 2006). Figure 2 (A) shows how ITL adapts to the smoothness of f: Under the "smooth" Gaussian kernel, points outside $\mathcal{A}$ provide higher-order information. In contrast, under the "rough" Laplace kernel and if $\mathcal{A} \subseteq \mathcal{S}$, points outside $\mathcal{A}$ do not provide any additional information, and therefore are not sampled by ITL. If, however, $\mathcal{A} \not\subseteq \mathcal{S}$, information "leaks" $\mathcal{A}$ even under a Laplace kernel prior. That is, even for non-smooth functions, the point with most information need not be in $\mathcal{A}$.

Does ITL outperform uncertainty sampling? Uncertainty sampling (UnSa, Lewis & Catlett, 1994) is one of the most popular active learning methods. UnSa selects points $\boldsymbol{x}$ with high *prior* uncertainty: $\boldsymbol{x}_n = \arg\max_{\boldsymbol{x} \in \mathcal{S}} \sigma^2_{n-1}(\boldsymbol{x})$. This is in stark contrast to ITL and VTL which select points $\boldsymbol{x}$ that minimize *posterior* (epistemic) uncertainty about $\mathcal{A}$. It can be seen that UnSa is the special "undirected" case of ITL when $\mathcal{S} \subseteq \mathcal{A}$ and observation noise is homoscedastic (cf. Appendix C.1).

We compare UnSa to ITL and VTL in Figure 2 (B). We observe that ITL and VTL outperform UnSa which also samples points that are not informative about $\mathcal{A}$. Further, ITL and VTL outperform "local" UnSa (i.e., UnSa constrained to $\mathcal{A} \cap \mathcal{S}$) which neglects all information provided by points outside $\mathcal{A}$.[4] As one would expect, VTL has an advantage with respect to reducing the total variance of $\boldsymbol{f}_{\mathcal{A}}$, whereas ITL reduces the entropy of $\boldsymbol{f}_{\mathcal{A}}$ faster. We include ablations in Appendix H where we, in particular, observe that the advantage of ITL and VTL over UnSa increases as the volume of prediction targets shrinks in comparison to the size of domain.

4 Safe Bayesian Optimization

Safe Bayesian optimization (Sui et al., 2015; Berkenkamp et al., 2021) has applications in natural science (Cooper & Netoff, 2022) and robotics (Wischnewski et al., 2019; Sukhija et al., 2023b;

[3] Here $f^\star(\boldsymbol{x})$ denotes the mean observation $y_{\boldsymbol{x}} = f^\star(\boldsymbol{x}) + \epsilon_{\boldsymbol{x}}$

[4] If $\mathcal{A} \not\subseteq \mathcal{S}$ then "local" UnSa does *not even* converge to the irreducible uncertainty.

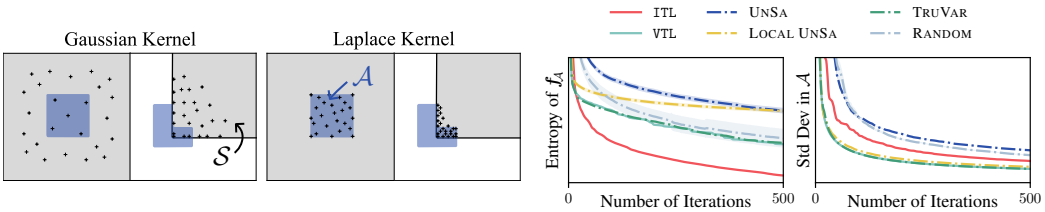

Figure 2: **(A)**: The left figure depicts initial 25 samples of ITL under a Gaussian kernel with lengthscale 1 (left) and a Laplace kernel with lengthscale 10 (right). Shown in gray is the sample space $\mathcal{S}$ and shown in blue is the target space $\mathcal{A}$. In three of the four examples, points outside the target space provide useful information. **(B)**: The right figure depicts the entropy of $\boldsymbol{f}_{\mathcal{A}}$ ranging from -3850 to -3725 and the mean marginal standard deviations of $\boldsymbol{f}_{\mathcal{A}}$ ranging from 0 to 0.15, under the Gaussian kernel in the left instance ($\mathcal{A} \subset \mathcal{S}$) from (A). It can be seen that ITL and VTL outperform UnSa and Random. Uncertainty bands correspond to one standard error over 10 random seeds.

Widmer et al., 2023). We show in this section that Safe BO can be cast as the "directed" learning problem introduced in Section 1. Safe BO solves the following optimization problem

$$\max_{\boldsymbol{x} \in \mathcal{S}^\star} f^\star(\boldsymbol{x}) \quad \text{where} \quad \mathcal{S}^\star = \{\boldsymbol{x} \in \mathcal{X} \mid g^\star(\boldsymbol{x}) \geq 0\} \tag{3}$$

which can be generalized to multiple constraints. The functions $f^\star$ and $g^\star$, and hence also the "safe set" $\mathcal{S}^\star$, are unknown and have to be actively learned from data. However, it is crucial that the data collection does not violate the constraint, i.e., $\boldsymbol{x}_n \in \mathcal{S}^\star, \forall n \geq 1$. In particular, Safe BO where the domain $\mathcal{X}$ comprises all action sequences of the length of one trajectory can be seen as an instance of episodic safe reinforcement learning.

Safe Bayesian optimization as Transductive Active Learning In the agnostic setting from Section 3.2, GPs f and g can be used as well-calibrated models of the ground truths $f^\star$ and $g^\star$, and we denote lower- and upper-confidence bounds by $l_n^f(\boldsymbol{x}), l_n^g(\boldsymbol{x})$ and $u_n^f(\boldsymbol{x}), u_n^g(\boldsymbol{x})$, respectively. These confidence bounds induce a *pessimistic* safe set $\mathcal{S}_n \stackrel{\text{def}}{=} \{\boldsymbol{x} \mid l_n^g(\boldsymbol{x}) \geq 0\}$ and an *optimistic* safe set $\widehat{\mathcal{S}}_n \stackrel{\text{def}}{=} \{\boldsymbol{x} \mid u_n^g(\boldsymbol{x}) \geq 0\}$ which satisfy $\mathcal{S}_n \subseteq \mathcal{S}^\star \subseteq \widehat{\mathcal{S}}_n$ with high probability at all times. Similarly, the set of *potential maximizers*

$$\mathcal{A}_n \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \widehat{\mathcal{S}}_n \mid u_n^f(\boldsymbol{x}) \geq \max_{\boldsymbol{x}' \in \mathcal{S}_n} l_n^f(\boldsymbol{x}')\} \tag{4}$$

contains the solution to Equation (3) at all times with high probability.

The (simple) regret $r_n(\mathcal{S}) \stackrel{\text{def}}{=} \max_{\boldsymbol{x} \in \mathcal{S}} f^\star(\boldsymbol{x}) - f^\star(\widehat{\boldsymbol{x}}_n)$ with $\widehat{\boldsymbol{x}}_n \stackrel{\text{def}}{=} \arg\max_{\boldsymbol{x} \in \mathcal{S}_n} l_n^f(\boldsymbol{x})$ measures the worst-case performance of a decision rule. To achieve small regret, one faces an *exploration-expansion* dilemma wherein one needs to explore points that are known-to-be-safe, i.e., lie in the estimated safe set $\mathcal{S}_n$, and might be optimal, while at the same time discovering new safe points by "expanding" $\mathcal{S}_n$. Accordingly, a natural choice for the target space of Safe BO is $\mathcal{A}_n$ since it captures both exploration and expansion *simultaneously*.[5] To prevent constraint violation, the sample space is restricted to the pessimistic safe set $\mathcal{S}_n$. In Safe BO, both the target space and sample space change with each round n, and we generalize our theoretical results from Section 3 in Appendix C to this setting.

Theorem 4.1 (Convergence to safe optimum)**.** *Pick any $\epsilon > 0$, $\delta \in (0, 1)$. Assume that $f^\star$, $g^\star$ lie in the reproducing kernel Hilbert space $\mathcal{H}_k(\mathcal{X})$ of the kernel k, and that the noise ε_n is conditionally ρ-sub-Gaussian. Then, we have with probability at least $1 - \delta$,*

Safety: for all $n \geq 1$, $\quad \boldsymbol{x}_n \in \mathcal{S}^\star$.

Moreover, assume $\mathcal{S}_0 \neq \emptyset$ and denote with $\mathcal{R}$ the largest reachable safe set starting from $\mathcal{S}_0$. Then, the convergence of reducible uncertainty implies that there exists $n^\star > 0$ such that with probability at least $1 - \delta$,

Optimality: for all $n \geq n^\star$, $\quad r_n(\mathcal{R}) \leq \epsilon$.

We provide a formal proof in Appendix C.7. Central to the proof is the application of Theorem 3.3 to show that the safety of parameters *outside* the safe set $\mathcal{S}_n$ can be inferred efficiently. In Section 3, we

[5] An alternative possibility is to weigh each point in $\mathcal{A}_n$ according to how likely it is to be the safe optimum. Which approach performs better is task-dependent, and we include a detailed discussion in Appendix I.1.

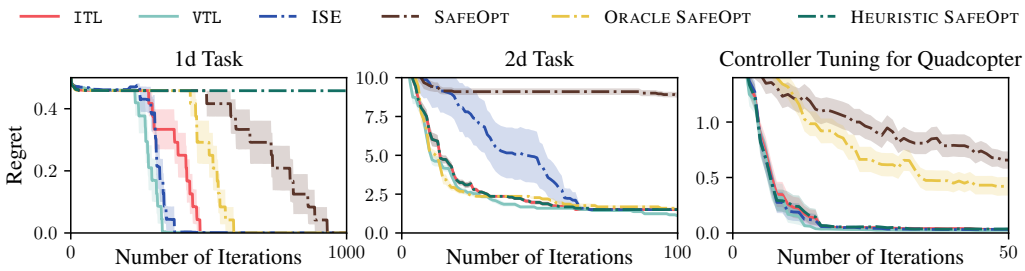

Figure 3: We compare ITL and VTL to Oracle SafeOpt, which has oracle knowledge of the Lipschitz constants, SafeOpt, where the Lipschitz constants are estimated from the GP, as well as Heuristic SafeOpt and ISE, and observe that ITL and VTL systematically perform well. We compare against additional baselines in Appendix I.1. The regret is evaluated with respect to the ground truth objective $f^\star$ and constraint $g^\star$, and averaged over 10 (in synthetic experiments) and 25 (in the quadcopter experiment) random seeds. Additional details can be found in Appendix I.4.

outline settings where the reducible uncertainty converges which is the case for a very general class of functions, and for such instances Theorem 4.1 guarantees optimality in the largest reachable safe set $\mathcal{R}$. $\mathcal{R}$ represents the largest set any safe learning algorithm can explore without violating the safety constraints (with high probability) during learning (cf. Definition C.29). Our guarantees are similar to those of other Safe BO algorithms (Berkenkamp et al., 2021) but require fewer assumptions and generalize to continuous domains. We obtain Theorem 4.1 from a more general result (Theorem C.34) which can be specialized to yield "free" novel convergence guarantees for problems other than Bayesian optimization, such as level set estimation, by choosing an appropriate target space.

4.1 Experiments on Safe Bayesian Optimization

We evaluate two synthetic experiments for a 1d and 2d parameter space, respectively (cf. Appendix I.4 for details), which demonstrate the various shortcomings of existing Safe BO baselines. Additionally, as third experiment, we safely tune the controller of a quadcopter.

Safe controller tuning for a quadcopter We consider a quadcopter with unknown dynamics; $\boldsymbol{s}_{t+1} = \boldsymbol{T}(\boldsymbol{s}_t, \boldsymbol{u}_t)$ where $\boldsymbol{u}_t \in \mathbb{R}^{d_u}$ is the control signal and $\boldsymbol{s}_t \in \mathbb{R}^{d_s}$ is the state at time t. The inputs $\boldsymbol{u}_t$ are calculated through a deterministic function of the state $\boldsymbol{\pi} : \mathcal{S} \to \mathcal{U}$ which we call the policy. The policy is parameterized via parameters $\boldsymbol{x} \in \mathcal{X}$, e.g., PID controller gains, such that $\boldsymbol{u}_t = \boldsymbol{\pi}_{\boldsymbol{x}}(\boldsymbol{s}_t)$. The goal is to find the optimal parameters with respect to an unknown objective $f^\star$ while satisfying some unknown constraint(s) $g^\star(\boldsymbol{x}) \geq 0$, e.g., the quadcopter does not fall on the ground. This is a typical Safe BO problem which is widely applied for safe controller learning in robotics (Berkenkamp et al., 2021; Baumann et al., 2021; Widmer et al., 2023).

Results We compare ITL and VTL to SafeOpt (Berkenkamp et al., 2021), which is undirected, i.e., expands in all directions including ones that are known-to-be suboptimal, and ISE (Bottero et al., 2022), which is solely expansionist — does not trade-off expansion-exploration. We provide a detailed discussion of baselines in Appendix I.2. In all our experiments, summarized in Figure 3, we observe that ITL and VTL systematically perform well, i.e., better or on par with the state-of-the-art. We attribute this to its directed exploration and less conservative expansion over SafeOpt (cf. 1d task and quadcopter experiment), and natural trade-off between expansion and exploration as opposed to ISE (see 2d task). Generally, VTL has a slight advantage over ITL, which is because VTL minimizes marginal variances (as opposed to entropy), which are decisive for expanding the safe set. While ITL and VTL do not violate constraints, we observe that other methods that do not explicitly enforce safety such as EIC (Gardner et al., 2014) lead to constraint violation (cf. Appendix I.4.2).

5 Related Work

Global "inductive" active learning Several works have previously found entropy-based decision rules to be useful for global active learning (Krause & Guestrin, 2007; Guo & Greiner, 2007; Krause et al., 2008) and semi-supervised learning (Grandvalet & Bengio, 2004). The variance-based VTL has previously been proposed by Cohn (1993) in the special case of global active learning without

proving theoretical guarantees. VTL was then recently re-derived by Shoham & Avron (2023) along other experimental design criteria under the lens of minimizing risk for global one-shot learning in overparameterized models. Substantial work on active learning has studied entropy-based criteria in *parameter-space*, most notably BALD (Houlsby et al., 2011; Gal et al., 2017; Kirsch et al., 2019), which selects $\boldsymbol{x}_n = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mathrm{I}(\boldsymbol{\theta}; y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1})$, where $\boldsymbol{\theta}$ is the random parameter vector of a parametric model (e.g., obtained via Bayesian deep learning). Such methods are inherently "undirected" in the sense that they do not facilitate learning on specific prediction targets.

Transductive active learning In contrast, ITL operates in *output-space* where it is straightforward to specify prediction targets, and which is computationally easier. Special cases of ITL when $\mathcal{S} = \mathcal{X}$ and $|\mathcal{A}| = 1$ have been proposed in the foundational work of MacKay (1992) on "directed" output-space active learning, and later studied empirically by Wang et al. (2021) and Smith et al. (2023). As generalization to larger target spaces, MacKay (1992) proposed mean-marginal ITL,

$$\boldsymbol{x}_n = \arg\max_{\boldsymbol{x}\in\mathcal{S}} \sum_{\boldsymbol{x}'\in\mathcal{A}} \mathrm{I}(f_{\boldsymbol{x}'}; y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1}) \,, \tag{MM-ITL}$$

for which we derive analogous versions of Theorems 3.2 and 3.3 in Appendix D.3. We note that similarly to VTL, MM-ITL disregards the mutual dependence of points in the target space $\mathcal{A}$ and differs from VTL only in a different weighting of the posterior marginal variances of the prediction targets (cf. Appendix D.3).

Other work on directed active learning Directed active learning methods have been proposed for the problem of determining the optimum of an unknown function, also known as best-arm identification (Audibert et al., 2010) or pure exploration bandits (Bubeck et al., 2009). Entropy search methods (Hennig & Schuler, 2012; Hernández-Lobato et al., 2014) are widely used and select $\boldsymbol{x}_n = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mathrm{I}(\boldsymbol{x}^*; y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1})$ in *input-space* where $\boldsymbol{x}^* = \arg\max_{\boldsymbol{x}} f_{\boldsymbol{x}}$. Similarly to ITL, *output-space* entropy search methods (Hoffman & Ghahramani, 2015; Wang & Jegelka, 2017), which select $\boldsymbol{x}_n = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mathrm{I}(f^*; y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1})$ with $f^* = \max_{\boldsymbol{x}} f_{\boldsymbol{x}}$, are more computationally tractable. In fact, output-space entropy search is a special case of ITL with a stochastic target space (cf. Equation (42) in Appendix I.1). Bogunovic et al. (2016) analyze TruVar in the context of Bayesian optimization and level set estimation. TruVar is akin to VTL with a similar notion of "target space", but their algorithm and analysis rely on a threshold scheme which requires that $\mathcal{A} \subseteq \mathcal{S}$ and is therefore not applicable to Safe BO. Fiez et al. (2019) introduce the *transductive linear bandit* problem, which is a special case of transductive active learning limited to a linear function class and with the objective of determining the maximum within an initial candidate set.[6] We mention additional more loosely related works in Appendix A.

6 Conclusion

We introduced transductive active learning, a generalization of classical active learning to concrete prediction targets and/or limited information due to constrained sample spaces. We showed that Safe BO can be formulated as a transductive active learning problem, and proposed ITL and VTL as simple and natural algorithms for this setting. We proved novel generalization bounds which we leveraged to show convergence guarantees that are stronger than those of prior works. Finally, we demonstrated empirically that ITL and VTL substantially outperform the state-of-the-art in terms of sample efficiency. We believe that transductive active learning is a promising lens that can also be applied to best policy identification in RL (e.g., Ménard et al., 2021) and other RL methods inspired by classical active learning (e.g., Sukhija et al., 2023a).

Acknowledgements

Many thanks to Armin Lederer, Johannes Kirschner, Jonas Rothfuss, Lars Lorch, Manish Prajapat, Nicolas Emmenegger, Parnian Kassraie, and Scott Sussex for their insightful feedback on different versions of this manuscript, as well as Anton Baumann for helpful discussions.

This project was supported in part by the European Research Council (ERC) under the European Union's Horizon 2020 research and Innovation Program Grant agreement no. 815943, the Swiss

[6] The transductive bandit problem can be solved analogously to Safe BO, by maintaining the set $\mathcal{A}_n$.

National Science Foundation under NCCR Automation, grant agreement 51NF40 180545, and by a grant of the Hasler foundation (grant no. 21039). Jonas Hübotter was supported in part by the German Academic Scholarship Foundation (Studienstiftung).

References

Abbasi-Yadkori, Y. *Online learning for linearly parametrized control problems*. PhD thesis, University of Alberta, 2013.

Audibert, J.-Y., Bubeck, S., and Munos, R. Best arm identification in multi-armed bandits. In *COLT*, 2010.

Balestriero, R., Ibrahim, M., Sobal, V., Morcos, A., Shekhar, S., Goldstein, T., Bordes, F., Bardes, A., Mialon, G., Tian, Y., et al. A cookbook of self-supervised learning. *arXiv preprint arXiv:2304.12210*, 2023.

Barrett, A. B. Exploration of synergistic and redundant information sharing in static and dynamical gaussian systems. *Physical Review E*, 91(5), 2015.

Baumann, D., Marco, A., Turchetta, M., and Trimpe, S. Gosafe: Globally optimal safe robot learning. In *ICRA*, 2021.

Bengio, Y., Louradour, J., Collobert, R., and Weston, J. Curriculum learning. In *ICML*, volume 26, 2009.

Beraha, M., Metelli, A. M., Papini, M., Tirinzoni, A., and Restelli, M. Feature selection via mutual information: New theoretical insights. In *IJCNN*, 2019.

Berkenkamp, F., Schoellig, A. P., and Krause, A. Safe controller optimization for quadrotors with gaussian processes. In *ICRA*, 2016.

Berkenkamp, F., Krause, A., and Schoellig, A. P. Bayesian optimization with safety constraints: safe and automatic parameter tuning in robotics. *Machine Learning*, 2021.

Bogunovic, I., Scarlett, J., Krause, A., and Cevher, V. Truncated variance reduction: A unified approach to bayesian optimization and level-set estimation. *NeurIPS*, 29, 2016.

Bottero, A., Luis, C., Vinogradska, J., Berkenkamp, F., and Peters, J. R. Information-theoretic safe exploration with gaussian processes. *NeurIPS*, 35, 2022.

Bottero, A. G., Luis, C. E., Vinogradska, J., Berkenkamp, F., and Peters, J. Information-theoretic safe bayesian optimization. *arXiv preprint arXiv:2402.15347*, 2024.

Bubeck, S., Munos, R., and Stoltz, G. Pure exploration in multi-armed bandits problems. In *ALT*, volume 20, 2009.

Chaloner, K. and Verdinelli, I. Bayesian experimental design: A review. *Statistical Science*, 1995.

Chandra, B. Quadrotor simulation, 2023. URL https://github.com/Bharath2/Quadrotor-Simulation.

Chowdhury, S. R. and Gopalan, A. On kernelized multi-armed bandits. In *ICML*, 2017.

Cohn, D. Neural network exploration using optimal experiment design. *NeurIPS*, 6, 1993.

Cooper, S. E. and Netoff, T. I. Multidimensional bayesian estimation for deep brain stimulation using the safeopt algorithm. *medRxiv*, 2022.

Cover, T. M. *Elements of information theory*. John Wiley & Sons, 1999.

Das, A. and Kempe, D. Algorithms for subset selection in linear regression. In *STOC*, volume 40, 2008.

Das, A. and Kempe, D. Approximate submodularity and its applications: Subset selection, sparse approximation and dictionary selection. *JMLR*, 19(1), 2018.

Emmenegger, N., Mutnỳ, M., and Krause, A. Likelihood ratio confidence sets for sequential decision making. *NeurIPS*, 37, 2023.

Fiez, T., Jain, L., Jamieson, K. G., and Ratliff, L. Sequential experimental design for transductive linear bandits. *NeurIPS*, 32, 2019.

Gal, Y., Islam, R., and Ghahramani, Z. Deep bayesian active learning with image data. In *ICML*, 2017.

Gao, M., Zhang, Z., Yu, G., Arık, S. Ö., Davis, L. S., and Pfister, T. Consistency-based semi-supervised active learning: Towards minimizing labeling cost. In *ECCV*, 2020.

Gardner, J. R., Kusner, M. J., Xu, Z. E., Weinberger, K. Q., and Cunningham, J. P. Bayesian optimization with inequality constraints. In *ICML*, volume 2014, 2014.

Grandvalet, Y. and Bengio, Y. Semi-supervised learning by entropy minimization. *NeurIPS*, 17, 2004.

Graves, A., Bellemare, M. G., Menick, J., Munos, R., and Kavukcuoglu, K. Automated curriculum learning for neural networks. In *ICML*, 2017.

Graybill, F. A. *An introduction to linear statistical models*. Literary Licensing, LLC, 1961.

Guo, Y. and Greiner, R. Optimistic active-learning using mutual information. In *IJCAI*, volume 7, 2007.

Hennig, P. and Schuler, C. J. Entropy search for information-efficient global optimization. *JMLR*, 13 (6), 2012.

Hernández-Lobato, J. M., Hoffman, M. W., and Ghahramani, Z. Predictive entropy search for efficient global optimization of black-box functions. *NeurIPS*, 27, 2014.

Hoffman, M. W. and Ghahramani, Z. Output-space predictive entropy search for flexible global optimization. In *NeurIPS workshop on Bayesian Optimization*, 2015.

Houlsby, N., Huszár, F., Ghahramani, Z., and Lengyel, M. Bayesian active learning for classification and preference learning. *CoRR*, 2011.

Kaddour, J., Sæmundsson, S., et al. Probabilistic active meta-learning. *NeurIPS*, 33, 2020.

Khanna, R., Elenberg, E., Dimakis, A., Negahban, S., and Ghosh, J. Scalable greedy feature selection via weak submodularity. In *AISTATS*, 2017.

Kirsch, A., Van Amersfoort, J., and Gal, Y. Batchbald: Efficient and diverse batch acquisition for deep bayesian active learning. *NeurIPS*, 32, 2019.

Kirschner, J., Mutny, M., Hiller, N., Ischebeck, R., and Krause, A. Adaptive and safe bayesian optimization in high dimensions via one-dimensional subspaces. In *ICML*, 2019.

Kothawade, S., Beck, N., Killamsetty, K., and Iyer, R. Similar: Submodular information measures based active learning in realistic scenarios. *NeurIPS*, 34, 2021.

Krause, A. and Golovin, D. Submodular function maximization. *Tractability*, 3, 2014.

Krause, A. and Guestrin, C. Nonmyopic active learning of gaussian processes: an exploration-exploitation approach. In *ICML*, volume 24, 2007.

Krause, A., Singh, A., and Guestrin, C. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *JMLR*, 9(2), 2008.

Lewis, D. D. and Catlett, J. Heterogeneous uncertainty sampling for supervised learning. In *Machine learning proceedings*. 1994.

MacKay, D. J. Information-based objective functions for §data selection. *Neural computation*, 4(4), 1992.

Ménard, P., Domingues, O. D., Jonsson, A., Kaufmann, E., Leurent, E., and Valko, M. Fast active learning for pure exploration in reinforcement learning. In *ICML*, 2021.

Murphy, K. P. *Probabilistic machine learning: Advanced topics*. MIT Press, 2023.

Mutny, M. and Krause, A. Experimental design for linear functionals in reproducing kernel hilbert spaces. *NeurIPS*, 35, 2022.

Nemhauser, G. L., Wolsey, L. A., and Fisher, M. L. An analysis of approximations for maximizing submodular set functions—i. *Mathematical programming*, 14, 1978.

Peng, H., Long, F., and Ding, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, 27(8), 2005.

Rothfuss, J., Koenig, C., Rupenyan, A., and Krause, A. Meta-learning priors for safe bayesian optimization. In *COLT*, 2023.

Russo, D. J., Van Roy, B., Kazerouni, A., Osband, I., Wen, Z., et al. A tutorial on thompson sampling. *Foundations and Trends® in Machine Learning*, 11(1), 2018.

Schreiter, J., Nguyen-Tuong, D., Eberts, M., Bischoff, B., Markert, H., and Toussaint, M. Safe exploration for active learning with gaussian processes. In *ECML PKDD*, 2015.

Seo, S., Wallat, M., Graepel, T., and Obermayer, K. Gaussian process regression: Active data selection and test point rejection. In *Mustererkennung 2000*. Springer, 2000.

Settles, B. Active learning literature survey. Technical report, University of Wisconsin-Madison Department of Computer Sciences, 2009.

Shoham, N. and Avron, H. Experimental design for overparameterized learning with application to single shot deep active learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

Shwartz-Ziv, R. and LeCun, Y. To compress or not to compress–self-supervised learning and information theory: A review. *arXiv preprint arXiv:2304.09355*, 2023.

Smith, F. B., Kirsch, A., Farquhar, S., Gal, Y., Foster, A., and Rainforth, T. Prediction-oriented bayesian active learning. In *AISTATS*, 2023.

Soviany, P., Ionescu, R. T., Rota, P., and Sebe, N. Curriculum learning: A survey. *IJCV*, 2022.

Srinivas, N., Krause, A., Kakade, S. M., and Seeger, M. Gaussian process optimization in the bandit setting: No regret and experimental design. In *ICML*, volume 27, 2009.

Sui, Y., Gotovos, A., Burdick, J., and Krause, A. Safe exploration for optimization with gaussian processes. In *ICML*, 2015.

Sukhija, B., Treven, L., Sancaktar, C., Blaes, S., Coros, S., and Krause, A. Optimistic active exploration of dynamical systems. *NeurIPS*, 37, 2023a.

Sukhija, B., Turchetta, M., Lindner, D., Krause, A., Trimpe, S., and Baumann, D. Gosafeopt: Scalable safe exploration for global optimization of dynamical systems. *Artificial Intelligence*, 2023b.

Thompson, W. R. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 1933.

Tu, S., Frostig, R., Singh, S., and Sindhwani, V. JAX: A python library for differentiable optimal control on accelerators, 2023. URL http://github.com/google/trajax.

Turchetta, M., Berkenkamp, F., and Krause, A. Safe exploration for interactive machine learning. *NeurIPS*, 32, 2019.

Vakili, S., Khezeli, K., and Picheny, V. On information gain and regret bounds in gaussian process bandits. In *AISTATS*, 2021.

Vapnik, V. *Estimation of dependences based on empirical data*. Springer Science & Business Media, 1982.

Vergara, J. R. and Estévez, P. A. A review of feature selection methods based on mutual information. *Neural computing and applications*, 24, 2014.

Wainwright, M. J. *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge university press, 2019.

Wang, C., Sun, S., and Grosse, R. Beyond marginal uncertainty: How accurately can bayesian regression models estimate posterior predictive correlations? In *AISTATS*, 2021.

Wang, Z. and Jegelka, S. Max-value entropy search for efficient bayesian optimization. In *ICML*, 2017.

Widmer, D., Kang, D., Sukhija, B., Hübotter, J., Krause, A., and Coros, S. Tuning legged locomotion controllers via safe bayesian optimization. *CORL*, 2023.

Wilks, S. S. Certain generalizations in the analysis of variance. *Biometrika*, 1932.

Williams, C. K. and Rasmussen, C. E. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA, 2006.

Wischnewski, A., Betz, J., and Lohmann, B. A model-free algorithm to safely approach the handling limit of an autonomous racecar. In *ICCVE*, 2019.

Yu, K., Bi, J., and Tresp, V. Active learning via transductive experimental design. In *ICML*, volume 23, 2006.

Zheng, H., Liu, R., Lai, F., and Prakash, A. Coverage-centric coreset selection for high pruning rates. *ICLR*, 2023.

Appendices

A general principle of "transductive learning" was already formulated by the famous computer scientist Vladimir Vapnik in the 20th century. Vapnik proposes the following "imperative for a complex world":

When solving a problem of interest, do not solve a more general problem as an intermediate step. Try to get the answer that you really need but not a more general one.

– Vapnik (1982)

These appendices provide additional background, proofs, experiment details, and ablation studies.

Contents

A Additional Related Work **14**

B Background **14**

B.1 Information Theory 14

B.2 Gaussian Processes 14

C Proofs **15**

C.1 Undirected Case of ITL 15

C.2 Non-adaptive Data Selection & Submodularity 15

C.3 Measures of Synergies & Approximate Submodularity 16

C.4 Convergence of Marginal Gain 18

C.5 Proof of Theorem 3.2 20

C.6 Proof of Theorem 3.3 25

C.7 Proof of Theorem 4.1 26

C.8 Useful Facts and Inequalities 30

D Interpretations & Approximations of Principle (†) **31**

D.1 Interpretations of ITL 31

D.2 Interpretations of VTL 31

D.3 Mean Marginal ITL 32

D.4 Summary 34

E Stochastic Target Spaces **34**

F Closed-form Decision Rules **35**

G Computational Complexity **35**

H Additional GP Experiments & Details **36**

I Additional Safe BO Experiments & Details **38**

I.1 A More Exploitative Stochastic Target Space 38

I.2 Detailed Comparison with Prior Works 39
I.3 Jumping Past Local Barriers 44
I.4 Experiment Details 45

A Additional Related Work

The general principle of non-active "transductive learning" was introduced by Vapnik (1982). The notion of "target" from transductive active learning is akin to the notion of "task" in curriculum learning (Bengio et al., 2009; Graves et al., 2017; Soviany et al., 2022). The study of settings where the irreducible uncertainty is zero is related to the study of estimability in experimental design (Graybill, 1961; Mutny & Krause, 2022). In feature selection, selecting features that maximize information gain with respect to a to-be-predicted label is a standard approach (Peng et al., 2005; Vergara & Estévez, 2014; Beraha et al., 2019) which is akin to ITL (cf. Appendix D). Balancing relevance and informativeness, similarly to ITL, is also important for data pruning (Zheng et al., 2023). Transductive active learning is complimentary to other learning methodologies, such as semi-supervised learning (Gao et al., 2020), self-supervised learning (Shwartz-Ziv & LeCun, 2023; Balestriero et al., 2023), and meta-learning (Kaddour et al., 2020; Rothfuss et al., 2023).

B Background

B.1 Information Theory

Throughout this work, $\log$ denotes the natural logarithm. Given random vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, we denote by

$$\begin{aligned} \mathrm{H}[\boldsymbol{x}] &\stackrel{\text{def}}{=} \mathbb{E}_{p(\boldsymbol{x})}[-\log p(\boldsymbol{x})], \\ \mathrm{H}[\boldsymbol{x} \mid \boldsymbol{y}] &\stackrel{\text{def}}{=} \mathbb{E}_{p(\boldsymbol{x},\boldsymbol{y})}[-\log p(\boldsymbol{x} \mid \boldsymbol{y})], \quad \text{and} \\ \mathrm{I}(\boldsymbol{x};\boldsymbol{y}) &\stackrel{\text{def}}{=} \mathrm{H}[\boldsymbol{x}] - \mathrm{H}[\boldsymbol{x} \mid \boldsymbol{y}] \end{aligned}$$

the (differential) entropy, conditional entropy, and information gain, respectively (Cover, 1999).[7]

The *multivariate information gain* (Murphy, 2023) between random vectors $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ is given by

$$\begin{aligned} \mathrm{I}(\boldsymbol{x};\boldsymbol{y};\boldsymbol{z}) &\stackrel{\text{def}}{=} \mathrm{I}(\boldsymbol{x};\boldsymbol{y}) - \mathrm{I}(\boldsymbol{x};\boldsymbol{y} \mid \boldsymbol{z}) && (5) \\ &= \mathrm{I}(\boldsymbol{x};\boldsymbol{y}) + \mathrm{I}(\boldsymbol{x};\boldsymbol{z}) - \mathrm{I}(\boldsymbol{x};\boldsymbol{y},\boldsymbol{z}). && (6) \end{aligned}$$

When $\mathrm{I}(\boldsymbol{x};\boldsymbol{y};\boldsymbol{z}) \neq 0$ it is said that $\boldsymbol{y}$ and $\boldsymbol{z}$ *interact* regarding their information about $\boldsymbol{x}$. If the interaction is positive, it is said that the information of $\boldsymbol{z}$ about $\boldsymbol{x}$ is *redundant* given $\boldsymbol{y}$. Conversely, if the interaction is negative, it is said that the information of $\boldsymbol{z}$ about $\boldsymbol{x}$ is *synergistic* with $\boldsymbol{y}$. The notion of synergy is akin to the frequentist notion of "suppressor variables" in linear regression (Das & Kempe, 2008).

B.2 Gaussian Processes

The stochastic process f is a Gaussian process (GP, Williams & Rasmussen (2006)), denoted $f \sim \mathcal{GP}(\mu, k)$, with mean function μ and kernel k if for any finite subset $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} \subseteq \mathcal{X}$, $\boldsymbol{f}_X \sim \mathcal{N}(\boldsymbol{\mu}_X, \boldsymbol{K}_{XX})$ is jointly Gaussian with mean vector $\boldsymbol{\mu}_X(i) = \mu(\boldsymbol{x}_i)$ and covariance matrix $\boldsymbol{K}_{XX}(i,j) = k(\boldsymbol{x}_i, \boldsymbol{x}_j)$.

In the following, we formalize the assumptions from the GP setting (cf. Section 3.1).

Assumption B.1 (Gaussian prior)**.** We assume that $f \sim \mathcal{GP}(\mu, k)$ with known mean function μ and kernel k.

Assumption B.2 (Gaussian noise)**.** We assume that the noise $\varepsilon_{\boldsymbol{x}}$ is mutually independent and zero-mean Gaussian with known variance $\rho^2(\boldsymbol{x}) > 0$. We write $\boldsymbol{P}_X = \operatorname{diag} \rho^2(\boldsymbol{x}_1), \ldots, \rho^2(\boldsymbol{x}_n)$.

Under Assumptions B.1 and B.2, the posterior distribution of f after observing points X is $\mathcal{GP}(\mu_n, k_n)$ with

$$\begin{aligned} \mu_n(\boldsymbol{x}) &= \mu(\boldsymbol{x}) + \boldsymbol{K}_{\boldsymbol{x}X}(\boldsymbol{K}_{XX} + \boldsymbol{P}_X)^{-1}(\boldsymbol{y}_X - \boldsymbol{\mu}_X), \\ k_n(\boldsymbol{x}, \boldsymbol{x'}) &= k(\boldsymbol{x}, \boldsymbol{x'}) - \boldsymbol{K}_{\boldsymbol{x}X}(\boldsymbol{K}_{XX} + \boldsymbol{P}_X)^{-1}\boldsymbol{K}_{X\boldsymbol{x'}}, \\ \sigma_n^2(\boldsymbol{x}) &= k_n(\boldsymbol{x}, \boldsymbol{x}). \end{aligned}$$

[7] One has to be careful to ensure that $\mathrm{I}(\boldsymbol{x};\boldsymbol{y})$ exists, i.e., $|\mathrm{I}(\boldsymbol{x};\boldsymbol{y})| < \infty$. We will assume that this is the case throughout this work. When $\boldsymbol{x}$ and $\boldsymbol{y}$ are jointly Gaussian, this is satisfied when the noise variance ρ^2 is positive.

For Gaussian random vectors $\boldsymbol{f}$ and $\boldsymbol{y}$, the entropy is $\mathrm{H}[\boldsymbol{f}] = \frac{n}{2}\log(2\pi e) + \frac{1}{2}\log|\mathrm{Var}[\boldsymbol{f}]|$, the information gain is $\mathrm{I}(\boldsymbol{f};\boldsymbol{y}) = \frac{1}{2}(\log|\mathrm{Var}[\boldsymbol{y}]| - \log|\mathrm{Var}[\boldsymbol{y} \mid \boldsymbol{f}]|)$, and

$$\gamma_n = \max_{\substack{X \subseteq \mathcal{X} \\ |X| \leq n}} \frac{1}{2}\log\left|\boldsymbol{I} + \boldsymbol{P}_X^{-1}\boldsymbol{K}_{XX}\right|.$$

C Proofs

We will write

- $\sigma^2 \stackrel{\text{def}}{=} \max_{\boldsymbol{x}\in\mathcal{X}} \sigma_0^2(\boldsymbol{x})$, and
- $\tilde{\sigma}^2 \stackrel{\text{def}}{=} \max_{\boldsymbol{x}\in\mathcal{X}} \sigma_0^2(\boldsymbol{x}) + \rho^2(\boldsymbol{x})$.

The following is a brief overview of the structure of this section:

1. Appendix C.1 relates ITL in the global learning setting ($\mathcal{S} \subseteq \mathcal{A}$) to prior work.
2. Appendix C.2 relates the designs selected by ITL and VTL to the optimal designs for corresponding non-adaptive objectives.
3. Appendix C.3 introduces measures of synergies that generalize the submodularity assumption (cf. Assumption 3.1).
4. Appendix C.4 proves key results on the convergence of the ITL and VTL objectives.
5. Appendix C.5 proves Theorem 3.2 (convergence in GP setting).
6. Appendix C.6 proves Theorem 3.3 (convergence in agnostic setting).
7. Appendix C.7 proves Theorem 4.1 (convergence in safe BO application).
8. Appendix C.8 includes useful facts.

C.1 Undirected Case of ITL

We briefly examine the important special case of ITL where $\mathcal{S} \subseteq \mathcal{A}$. In this setting, for all $\boldsymbol{x} \in \mathcal{S}$, the decision rule of ITL simplifies to

$$\begin{aligned}
\mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}} \mid \mathcal{D}_n) &\stackrel{(i)}{=} \mathrm{I}(\boldsymbol{f}_{\mathcal{A}\setminus\{\boldsymbol{x}\}}; y_{\boldsymbol{x}} \mid f_{\boldsymbol{x}}, \mathcal{D}_n) + \mathrm{I}(f_{\boldsymbol{x}}; y_{\boldsymbol{x}} \mid \mathcal{D}_n) \\
&\stackrel{(ii)}{=} \mathrm{I}(f_{\boldsymbol{x}}; y_{\boldsymbol{x}} \mid \mathcal{D}_n) \\
&= \mathrm{H}[y_{\boldsymbol{x}} \mid \mathcal{D}_n] - \mathrm{H}[\varepsilon_{\boldsymbol{x}}]
\end{aligned}$$

where (i) follows from the chain rule of information gain and $\boldsymbol{x} \in \mathcal{S} \subseteq \mathcal{A}$; and (ii) follows from the conditional independence $\boldsymbol{f}_{\mathcal{A}} \perp y_{\boldsymbol{x}} \mid f_{\boldsymbol{x}}$.

If additionally f is a GP then

$$\mathrm{H}[y_{\boldsymbol{x}} \mid \mathcal{D}_n] - \mathrm{H}[\varepsilon_{\boldsymbol{x}}] = \frac{1}{2}\log\left(1 + \frac{\mathrm{Var}[f_{\boldsymbol{x}} \mid \mathcal{D}_n]}{\mathrm{Var}[\varepsilon_{\boldsymbol{x}}]}\right).$$

This decision rule has also been termed *total information gain* (MacKay, 1992). When $\mathcal{S} \subseteq \mathcal{A}$ and observation noise is homoscedastic, this decision rule is equivalent to uncertainty sampling.

C.2 Non-adaptive Data Selection & Submodularity

Recall the non-myopic information gain $\psi_{\mathcal{A}}(X) = \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; \boldsymbol{y}_X)$ (ITL) and variance reduction $\psi_{\mathcal{A}}(X) = \mathrm{tr}\ \mathrm{Var}[\boldsymbol{f}_{\mathcal{A}}] - \mathrm{tr}\ \mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_X]$ (VTL) objective functions from Assumption 3.1. In this section, we will relate the designs selected by ITL and VTL to the optimal designs for these objectives. To this end, consider the non-adaptive optimization problem

$$X^\star = \arg\max_{\substack{X \subseteq \mathcal{S} \\ |X| = k}} \psi_{\mathcal{A}}(X).$$

Lemma C.1. *For both* ITL *and* VTL*,* $\psi_{\mathcal{A}}$ *is non-negative and monotone.*

Proof. For ITL, $\psi_{\mathcal{A}}(X) \geq 0$ follows from the non-negativity of mutual information. To conclude monotonicity, note that for any $X' \subseteq X \subseteq \mathcal{S}$,

$$\mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; \boldsymbol{y}_{X'}) = \mathrm{H}[\boldsymbol{f}_{\mathcal{A}}] - \mathrm{H}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_{X'}] \leq \mathrm{H}[\boldsymbol{f}_{\mathcal{A}}] - \mathrm{H}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_{X}] = \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; \boldsymbol{y}_{X})$$

due to monotonicity of conditional entropy (which is also called the "information never hurts" principle).

For VTL, recall that $\mathrm{tr}\ \mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_X] \leq \mathrm{tr}\ \mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_{X'}]$ for any $X' \subseteq X \subseteq \mathcal{S}$ (with an implicit expectation over $\boldsymbol{y}_X, \boldsymbol{y}_{X'}$). Non-negativity and monotonicity of $\psi_{\mathcal{A}}$ then follow analogously to ITL. □

Lemma C.2. *The marginal gain* $\Delta_{\mathcal{A}}(\boldsymbol{x} \mid X) \stackrel{\text{def}}{=} \psi_{\mathcal{A}}(X \cup \{\boldsymbol{x}\}) - \psi_{\mathcal{A}}(X)$ *of* $\boldsymbol{x} \in \mathcal{S}$ *given* $X \subseteq \mathcal{S}$ *is the* ITL *and* VTL *objective, respectively.*

Proof. For ITL,

$$\begin{aligned}\Delta_{\mathcal{A}}(\boldsymbol{x} \mid X) &= \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; \boldsymbol{y}_X, y_{\boldsymbol{x}}) - \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; \boldsymbol{y}_X) \\ &= \mathrm{H}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_X] - \mathrm{H}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_X, y_{\boldsymbol{x}}] \\ &= \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}} \mid \boldsymbol{y}_X)\end{aligned}$$

which is precisely the ITL objective.

For VTL,

$$\begin{aligned}\Delta_{\mathcal{A}}(\boldsymbol{x} \mid X) &= \mathrm{tr}\ \mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_X] - \mathrm{tr}\ \mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_X, y_{\boldsymbol{x}}] \\ &= -\mathrm{tr}\ \mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_X, y_{\boldsymbol{x}}] + \mathrm{const}\end{aligned}$$

which is precisely the VTL objective. □

Definition C.3 (Submodularity)**.** $\psi_{\mathcal{A}}$ is submodular if and only if for all $\boldsymbol{x} \in \mathcal{S}$ and $X' \subseteq X \subseteq \mathcal{S}$,

$$\Delta_{\mathcal{A}}(\boldsymbol{x} \mid X') \geq \Delta_{\mathcal{A}}(\boldsymbol{x} \mid X).$$

Theorem C.4 (Nemhauser et al. (1978))**.** *Let Assumption 3.1 hold. For any* $n \geq 1$*, if* ITL *or* VTL *selected* $\boldsymbol{x}_{1:n}$*, respectively, then*

$$\psi_{\mathcal{A}}(\boldsymbol{x}_{1:n}) \geq (1 - 1/e) \max_{\substack{X \subseteq \mathcal{S} \\ |X| \leq n}} \psi_{\mathcal{A}}(X).$$

Proof. This is a special case of a canonical result from non-negative monotone submodular function maximization (Nemhauser et al., 1978; Krause & Golovin, 2014). □

C.3 Measures of Synergies & Approximate Submodularity

We will now show that "downstream synergies", if present, can be seen as a source of learning complexity, which is orthogonal to the information capacity γ_n.

Example C.5. Consider the example where f is a stochastic process of three random variables X, Y, Z where X and Y are Bernoulli ($p = \frac{1}{2}$), and Z is the XOR of X and Y. Suppose that observations are exact (i.e., $\varepsilon_n = 0$), that the target space $\mathcal{A}$ comprises the output variable Z while the sample space $\mathcal{S}$ comprises the input variables X and Y. Observing any single X or Y yields no information about Z: $\mathrm{I}(Z; X) = \mathrm{I}(Z; Y) = 0$, however, observing both inputs jointly perfectly determines Z: $\mathrm{I}(Z; X, Y) = 1$. Thus, $\gamma_n(\mathcal{A}; \mathcal{S}) = 1$ if $n \geq 2$ and $\gamma_n(\mathcal{A}; \mathcal{S}) = 0$ else.

Learning about Z in examples of this kind is difficult for agents that make decisions greedily, since the next action (observing X or Y) yields no signal about its long-term usefulness. We call a sequence of observations, such as $\{X, Y\}$, *synergistic* since its combined information value is larger than the individual values. The prevalence of synergies is not captured by the information capacity $\gamma_n(\mathcal{A}; \mathcal{S})$ since it measures only the joint information gain of n samples within $\mathcal{S}$. Instead, the prevalence of synergies is captured by the sequence $\Gamma_n \stackrel{\text{def}}{=} \max_{\boldsymbol{x} \in \mathcal{S}} \Delta_{\mathcal{A}}(\boldsymbol{x} \mid \boldsymbol{x}_{1:n})$, which measures the maximum information gain of y_{n+1}. If $\Gamma_n > \Gamma_{n-1}$ at any round n, this indicates a synergy. The following key object measures the additional complexity due to synergies.

Definition C.6 (Task complexity)**.** For $n \geq 1$, assuming $\Gamma_i > 0$ for all $1 \leq i \leq n$, we define the *task complexity* as

$$\alpha_{\mathcal{A},\mathcal{S}}(n) \stackrel{\text{def}}{=} \max_{i \in \{0, \ldots, n-1\}} \frac{\Gamma_{n-1}}{\Gamma_i}.$$

Note that $\alpha_{\mathcal{A},\mathcal{S}}(n)$ is large only if the information gain of y_n is larger than that of a previous observation y_i. Intuitively, if $\alpha_{\mathcal{A},\mathcal{S}}(n)$ is large, the agent had to discover the *implicit* intermediate observations $y_1, \dots, y_{n-1}$ that lead to downstream synergies. We will subsequently formalize the intimate connections of the task complexity to synergies and submodularity. Note that in the GP setting, $\alpha_{\mathcal{A},\mathcal{S}}(n)$ can be computed online by keeping track of the smallest Γ_i during previous rounds i. Further, note that $\alpha_{\mathcal{A},\mathcal{S}}(n) \leq 1$ if $\psi_{\mathcal{A}}$ is submodular.

C.3.1 The Information Ratio

Another object will prove useful in our analysis of synergies.

Consider an alternative multiplicative interpretation of the multivariate information gain (cf. Equation (6)), which we call the *information ratio* of $X \subseteq \mathcal{S}$ given $D \subseteq \mathcal{S}$, $|X|, |D| < \infty$:

$$\bar{\kappa}(X \mid D) \stackrel{\text{def}}{=} \frac{\sum_{\boldsymbol{x} \in X} \Delta_{\mathcal{A}}(\boldsymbol{x} \mid D)}{\Delta_{\mathcal{A}}(X \mid D)} \in [0, \infty). \tag{7}$$

Observe that $\bar{\kappa}(X \mid D)$ measures the synergy properties of $\boldsymbol{y}_X$ with respect to $\boldsymbol{f}_{\mathcal{A}}$ given $\boldsymbol{y}_D$ in a multiplicative sense. That is, if $\bar{\kappa}(X \mid D) > 1$ then information in $\boldsymbol{y}_X$ is redundant, whereas if $\bar{\kappa}(X \mid D) < 1$ then information in $\boldsymbol{y}_X$ is synergistic, and if $\bar{\kappa}(X \mid D) = 1$ then $\boldsymbol{y}_X$ do not mutually interact with respect to $\boldsymbol{f}_{\mathcal{A}}$ (all given $\boldsymbol{y}_D$). In the degenerate case where $\Delta_{\mathcal{A}}(X \mid D) = 0$ (which implies $\sum_{\boldsymbol{x} \in X} \Delta_{\mathcal{A}}(\boldsymbol{x} \mid D) = 0$) we therefore let $\bar{\kappa}(X \mid D) = 1$.

The information ratio of ITL is strictly positive in the Gaussian case We prove the following straightforward lower bound to the information ratio of ITL.

Lemma C.7. *Let $X, D \subseteq \mathcal{S}, |X|, |D| < \infty$. If $\boldsymbol{f}_{\mathcal{A}}$ and $\boldsymbol{y}_{X \cup D}$ are jointly Gaussian then $\bar{\kappa}(X \mid D) > 0$.*

Proof. W.l.o.g. assume $D = \emptyset$. We let $X = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_k\}$ and prove lower and upper bound separately. We assume w.l.o.g. that $\mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; \boldsymbol{y}_X) > 0$ which implies $|\mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_X]| < |\mathrm{Var}[\boldsymbol{f}_{\mathcal{A}}]|$. Thus, there exists some i such that $\boldsymbol{f}_{\mathcal{A}}$ and $y_{\boldsymbol{x}_i}$ are dependent, so $|\mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid y_{\boldsymbol{x}_i}]| < |\mathrm{Var}[\boldsymbol{f}_{\mathcal{A}}]|$ which implies $\mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}_i}) > 0$. We therefore conclude that $\bar{\kappa}(X) > 0$. □

The following example shows that this lower bound is tight.

Example C.8 (Synergies of Gaussian random variables, inspired by Section 3 of Barrett (2015))**.** Consider the three random variables X, Y, and Z (think $\mathcal{A} = \{X\}$ and $\mathcal{S} = \{Y, Z\}$) which are jointly Gaussian with mean vector $\mathbf{0}$ and covariance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & a & a \\ a & 1 & 0 \\ a & 0 & 1 \end{bmatrix}, \qquad \text{for } 2a^2 < 1$$

where the constraint on a is to ensure that $\boldsymbol{\Sigma}$ is positive definite. Computing the mutual information, we have

$$\mathrm{I}(X; Y) = \mathrm{I}(X; Z) = -\frac{1}{2} \log(1 - a^2)$$

and $\mathrm{I}(X; Y, Z) = -\frac{1}{2} \log(1 - 2a^2)$. Therefore,

$$\frac{\mathrm{I}(X; Y) + \mathrm{I}(X; Z)}{\mathrm{I}(X; Y, Z)} = \frac{\log(1 - 2a^2 + a^4)}{\log(1 - 2a^2)} < 1.$$

Note that

$$\lim_{a \to \frac{1}{\sqrt{2}}} \frac{\log(1 - 2a^2 + a^4)}{\log(1 - 2a^2)} = 0,$$

and hence — perhaps unintuitively — even if Y and Z are uncorrelated, their information about X may be arbitrarily synergistic.

C.3.2 The Submodularity of the Special "Undirected" Case of ITL

In the global transductive active learning problem considered in most prior works, where $\mathcal{S} \subseteq \mathcal{A}$ and f is a Gaussian process, it holds for ITL that $\alpha_{\mathcal{A},\mathcal{S}}(n) = 1$ since all learning targets appear *explicitly* in $\mathcal{S}$:

Lemma C.9. *Let $\mathcal{S} \subseteq \mathcal{A}$. Then $\psi_{\mathcal{A}}$ of* ITL *is submodular.*

Proof. Fix any $\boldsymbol{x} \in \mathcal{S}$ and $X' \subseteq X \subseteq \mathcal{S}$. Let $\bar{X} \overset{\text{def}}{=} X \setminus X'$. By the definition of conditional information gain, we have

$$\Delta_{\mathcal{A}}(\boldsymbol{x} \mid X) = \mathrm{I}(y_{\boldsymbol{x}}; \boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_X) = \mathrm{I}(y_{\boldsymbol{x}}; \boldsymbol{f}_{\mathcal{A}}, \boldsymbol{y}_{X'} \mid \boldsymbol{y}_{\bar{X}}) - \mathrm{I}(y_{\boldsymbol{x}}; \boldsymbol{y}_{X'} \mid \boldsymbol{y}_{\bar{X}}).$$

Since for any $\boldsymbol{x} \in \mathcal{S}$ and $X \subseteq \mathcal{S}$, $y_{\boldsymbol{x}} \perp \boldsymbol{y}_X \mid \boldsymbol{f}_{\mathcal{A}}$, this simplifies to

$$\mathrm{I}(y_{\boldsymbol{x}}; \boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_X) = \mathrm{I}(y_{\boldsymbol{x}}; \boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_{\bar{X}}) - \mathrm{I}(y_{\boldsymbol{x}}; \boldsymbol{y}_{X'} \mid \boldsymbol{y}_{\bar{X}}).$$

It then follows from $\mathrm{I}(y_{\boldsymbol{x}}; \boldsymbol{y}_{X'} \mid \boldsymbol{y}_{\bar{X}}) \geq 0$ that

$$\Delta_{\mathcal{A}}(\boldsymbol{x} \mid X) = \mathrm{I}(y_{\boldsymbol{x}}; \boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_X) \leq \mathrm{I}(y_{\boldsymbol{x}}; \boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_{\bar{X}}) = \Delta_{\mathcal{A}}(\boldsymbol{x} \mid X').$$

□

This implies that $\alpha_{\mathcal{A},\mathcal{S}}(n) \leq 1$ for any n and $\bar{\kappa}(X \mid D) \geq 1$ for any $X, D \subseteq \mathcal{S}$ when $\mathcal{S} \subseteq \mathcal{A}$.

C.3.3 The Submodularity Ratio

Building upon the theory of maximizing non-negative monotone submodular functions (Nemhauser et al., 1978; Krause & Golovin, 2014), Das & Kempe (2018) define the following notion of "approximate" submodularity:

Definition C.10 (Submodularity ratio)**.** The *submodularity ratio* of $\psi_{\mathcal{A}}$ up to cardinality $n \geq 1$ is

$$\kappa_{\mathcal{A}}(n) \overset{\text{def}}{=} \min_{\substack{D \subseteq \boldsymbol{x}_{1:n} \\ X \subseteq \mathcal{S} : |X| \leq n \\ D \cap X = \emptyset}} \bar{\kappa}(X \mid D), \tag{8}$$

where they define $\frac{0}{0} \equiv 1$. $\psi_{\mathcal{A}}$ is said to be *κ-weakly submodular* for some $\kappa > 0$ if $\inf_{n \in \mathbb{N}} \kappa_{\mathcal{A}}(n) \geq \kappa$.

As a special case of Theorem 6 from Das & Kempe (2018), applying that $\psi_{\mathcal{A}}$ is non-negative and monotone, we obtain the following result.

Theorem C.11 (Das & Kempe (2018))**.** *For any $n \geq 1$, if* ITL *or* VTL *selected $\boldsymbol{x}_{1:n}$, respectively, then*

$$\psi_{\mathcal{A}}(\boldsymbol{x}_{1:n}) \geq (1 - e^{-\kappa_{\mathcal{A}}(n)}) \max_{\substack{X \subseteq \mathcal{S} \\ |X| \leq n}} \psi_{\mathcal{A}}(X).$$

If $\psi_{\mathcal{A}}$ is submodular, it is implied that $\kappa_{\mathcal{A}}(n) \geq 1$ for all $n \geq 1$ in which case Theorem C.11 recovers Theorem C.4.

C.4 Convergence of Marginal Gain

Our following analysis allows for changing target spaces $\mathcal{A}_n$ and sample spaces $\mathcal{S}_n$ (cf. Section 4), and to this end, we redefine $\Gamma_n \overset{\text{def}}{=} \max_{\boldsymbol{x} \in \mathcal{S}_n} \Delta_{\mathcal{A}_n}(\boldsymbol{x} \mid \boldsymbol{x}_{1:n})$. The following theorems show that the marginal gains of ITL and VTL converge to zero, and will serve as the main tool for establishing Theorems 3.2 and 3.3. We will abbreviate $\alpha_{\mathcal{A},\mathcal{S}}(n)$ by α_n.

Theorem C.12 (Convergence of Marginal Gain for ITL)**.** *Assume that Assumptions B.1 and B.2 are satisfied. Fix any integers $n_1 > n_0 \geq 0$, $\Delta = n_1 - n_0 + 1$ such that for all $i \in \{n_0, \ldots, n_1 - 1\}$, $\mathcal{A}_{i+1} \subseteq \mathcal{A}_i$ and $\mathcal{S} \overset{\text{def}}{=} \mathcal{S}_{i+1} = \mathcal{S}_i$. Further, assume $|\mathcal{A}_{n_0}| < \infty$. Then, if the sequence $\{\boldsymbol{x}_{i+1}\}_{i=n_0}^{n_1}$ was generated by* ITL*,*

$$\Gamma_{n_1} \leq \alpha_{n_1} \frac{\gamma_\Delta}{\Delta}. \tag{9}$$

Moreover, if $n_0 = 0$,

$$\Gamma_{n_1} \leq \alpha_{n_1} \frac{\gamma_{\mathcal{A}_0,\mathcal{S}}(\Delta)}{\Delta}. \tag{10}$$

Proof. We have

$$
\begin{aligned}
\Gamma_{n_1} &= \frac{1}{\Delta}\sum_{i=n_0}^{n_1}\Gamma_{n_1} \\
&\overset{(i)}{\leq} \frac{\alpha_{n_1}}{\Delta}\sum_{i=n_0}^{n_1}\Gamma_i \\
&= \frac{\alpha_{n_1}}{\Delta}\sum_{i=n_0}^{n_1}\max_{\boldsymbol{x}\in\mathcal{S}}\mathrm{I}(\boldsymbol{f}_{\mathcal{A}_i};y_{\boldsymbol{x}}\mid \boldsymbol{y}_{1:i}) \\
&\overset{(ii)}{=} \frac{\alpha_{n_1}}{\Delta}\sum_{i=n_0}^{n_1}\mathrm{I}(\boldsymbol{f}_{\mathcal{A}_i};y_{\boldsymbol{x}_{i+1}}\mid \mathcal{D}_i) \\
&\overset{(iii)}{\leq} \frac{\alpha_{n_1}}{\Delta}\sum_{i=n_0}^{n_1}\mathrm{I}(\boldsymbol{f}_{\mathcal{A}_{n_0}};y_{\boldsymbol{x}_{i+1}}\mid \mathcal{D}_i) \\
&\overset{(iv)}{=} \frac{\alpha_{n_1}}{\Delta}\sum_{i=n_0}^{n_1}\mathrm{I}(\boldsymbol{f}_{\mathcal{A}_{n_0}};y_{\boldsymbol{x}_{i+1}}\mid \boldsymbol{y}_{\boldsymbol{x}_{n_0+1:i}},\mathcal{D}_{n_0}) \\
&\overset{(v)}{=} \frac{\alpha_{n_1}}{\Delta}\mathrm{I}(\boldsymbol{f}_{\mathcal{A}_{n_0}};\boldsymbol{y}_{\boldsymbol{x}_{n_0+1:n_1+1}}\mid \mathcal{D}_{n_0}) \\
&\leq \frac{\alpha_{n_1}}{\Delta}\max_{\substack{X\subseteq\mathcal{S}\\|X|=\Delta}}\mathrm{I}(\boldsymbol{f}_{\mathcal{A}_{n_0}};\boldsymbol{y}_X\mid \mathcal{D}_{n_0}) \\
&\overset{(vi)}{\leq} \frac{\alpha_{n_1}}{\Delta}\max_{\substack{X\subseteq\mathcal{S}\\|X|=\Delta}}\mathrm{I}(\boldsymbol{f}_X;\boldsymbol{y}_X\mid \mathcal{D}_{n_0}) \\
&\overset{(vii)}{\leq} \frac{\alpha_{n_1}}{\Delta}\max_{\substack{X\subseteq\mathcal{S}\\|X|=\Delta}}\mathrm{I}(\boldsymbol{f}_X;\boldsymbol{y}_X) \\
&= \alpha_{n_1}\frac{\gamma_\Delta}{\Delta}
\end{aligned}
$$

where (i) follows from the definition of the task complexity α_{n_1} (cf. Definition C.6); (ii) uses the objective of ITL and that the posterior variance of Gaussians is independent of the realization and only depends on the *location* of observations; (iii) uses $\mathcal{A}_{i+1}\subseteq\mathcal{A}_i$ and monotonicity of information gain; (iv) uses that the posterior variance of Gaussians is independent of the realization and only depends on the *location* of observations; (v) uses the chain rule of information gain; (vi) uses $\boldsymbol{y}_X \perp \boldsymbol{f}_{\mathcal{A}_{n_0}} \mid \boldsymbol{f}_X$ and the data processing inequality. The conditional independence follows from the assumption that the observation noise is independent. Similarly, $\boldsymbol{y}_X \perp \mathcal{D}_{n_0} \mid \boldsymbol{f}_X$ which implies (vii).

If $n_0 = 0$, then the bound before line (vi) simplifies to $\alpha_{n_1}\gamma_{\mathcal{A}_0,\mathcal{S}}(\Delta)/\Delta$. □

The result for VTL is stated, for simplicity, only for the case where the target space and sample space are fixed.

Theorem C.13 (Convergence of Marginal Gain for VTL)**.** *Assume that Assumptions B.1 and B.2 are satisfied. Then for any $n \geq 1$, if the sequence $\{\boldsymbol{x}_i\}_{i=1}^n$ is generated by* VTL*,*

$$
\Gamma_{n-1} \leq \frac{2\sigma^2\alpha_n}{n}\sum_{\boldsymbol{x}'\in\mathcal{A}}\gamma_{\{\boldsymbol{x}'\},\mathcal{S}}(n). \tag{11}
$$

We remark that $\sum_{\boldsymbol{x}'\in\mathcal{A}}\gamma_{\{\boldsymbol{x}'\},\mathcal{S}}(n) \leq |\mathcal{A}|\gamma_{\mathcal{A},\mathcal{S}}(n)$.

Proof. We have

$$
\Gamma_{n-1} = \frac{1}{n}\sum_{i=0}^{n-1}\Gamma_{n-1}
$$

$$
\begin{aligned}
&\overset{(i)}{\leq} \frac{\alpha_n}{n} \sum_{i=0}^{n-1} \Gamma_i \\
&= \frac{\alpha_n}{n} \sum_{i=0}^{n-1} \Big[\operatorname{tr} \operatorname{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_{1:i}] - \min_{\boldsymbol{x} \in \mathcal{S}} \operatorname{tr} \operatorname{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_{1:i}, y_{\boldsymbol{x}}] \Big] \\
&\overset{(ii)}{=} \frac{\alpha_n}{n} \sum_{i=0}^{n-1} \Big[\operatorname{tr} \operatorname{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_i] - \operatorname{tr} \operatorname{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_{i+1}] \Big] \\
&\overset{(iii)}{\leq} \frac{\sigma^2 \alpha_n}{n} \sum_{\boldsymbol{x}' \in \mathcal{A}} \sum_{i=0}^{n-1} \log \left(\frac{\operatorname{Var}[f_{\boldsymbol{x}'} \mid \mathcal{D}_n]}{\operatorname{Var}[f_{\boldsymbol{x}'} \mid \mathcal{D}_{n+1}]} \right) \\
&= \frac{2\sigma^2 \alpha_n}{n} \sum_{\boldsymbol{x}' \in \mathcal{A}} \sum_{i=0}^{n-1} \mathrm{I}\big(f_{\boldsymbol{x}'}; y_{\boldsymbol{x}_{n+1}} \,\big|\, \mathcal{D}_n\big) \\
&\overset{(iv)}{=} \frac{2\sigma^2 \alpha_n}{n} \sum_{\boldsymbol{x}' \in \mathcal{A}} \sum_{i=0}^{n-1} \mathrm{I}\big(f_{\boldsymbol{x}'}; y_{\boldsymbol{x}_{n+1}} \,\big|\, \boldsymbol{y}_{\boldsymbol{x}_{1:n}}\big) \\
&\overset{(v)}{=} \frac{2\sigma^2 \alpha_n}{n} \sum_{\boldsymbol{x}' \in \mathcal{A}} \mathrm{I}\big(f_{\boldsymbol{x}'}; \boldsymbol{y}_{\boldsymbol{x}_{1:n}}\big) \\
&\leq \frac{2\sigma^2 \alpha_n}{n} \sum_{\boldsymbol{x}' \in \mathcal{A}} \max_{\substack{X \subseteq \mathcal{S} \\ |X| = n}} \mathrm{I}(f_{\boldsymbol{x}'}; \boldsymbol{y}_X) \\
&= \frac{2\sigma^2 \alpha_n}{n} \sum_{\boldsymbol{x}' \in \mathcal{A}} \gamma_{\{\boldsymbol{x}'\}, \mathcal{S}}(n)
\end{aligned}
$$

where (i) follows from the definition of the task complexity α_{n_1} (cf. Definition C.6); (ii) follows from the VTL decision rule and that the posterior variance of Gaussians is independent of the realization and only depends on the *location* of observations; (iii) follows from Lemma C.38 and monotonicity of variance; (iv) uses that the posterior variance of Gaussians is independent of the realization and only depends on the *location* of observations; and (v) uses the chain rule of mutual information. The remainder of the proof is analogous to the proof of Theorem C.12 (cf. Appendix C.4). □

Keeping track of the task complexity online In general, the task complexity α_n may be larger than one in the "directed" setting (i.e., when $\mathcal{S} \not\subseteq \mathcal{A}$). However, note that α_n can easily be evaluated online by keeping track of the smallest Γ_i during previous rounds i.

C.5 Proof of Theorem 3.2

We will now prove Theorem 3.2. We first prove the convergence of marginal variance within $\mathcal{S}$ for ITL, before proving the convergence outside $\mathcal{S}$ in Appendix C.5.1.

Lemma C.14 (Uniform convergence of marginal variance within $\mathcal{S}$ for ITL)**.** *Assume that Assumptions B.1 and B.2 are satisfied. For any $n \geq 0$ and $\boldsymbol{x} \in \mathcal{A} \cap \mathcal{S}$,*

$$
\sigma_n^2(\boldsymbol{x}) \leq 2\tilde{\sigma}^2 \cdot \Gamma_n. \tag{12}
$$

Proof. We have

$$
\begin{aligned}
\sigma_n^2(\boldsymbol{x}) &= \operatorname{Var}[f_{\boldsymbol{x}} \mid \mathcal{D}_n] - \underbrace{\operatorname{Var}[f_{\boldsymbol{x}} \mid f_{\boldsymbol{x}}, \mathcal{D}_n]}_{0} \\
&\overset{(i)}{=} \operatorname{Var}[y_{\boldsymbol{x}} \mid \mathcal{D}_n] - \rho^2(\boldsymbol{x}) - (\operatorname{Var}[y_{\boldsymbol{x}} \mid f_{\boldsymbol{x}}, \mathcal{D}_n] - \rho^2(\boldsymbol{x})) \\
&= \operatorname{Var}[y_{\boldsymbol{x}} \mid \mathcal{D}_n] - \operatorname{Var}[y_{\boldsymbol{x}} \mid f_{\boldsymbol{x}}, \mathcal{D}_n] \\
&\overset{(ii)}{\leq} \tilde{\sigma}^2 \log \left(\frac{\operatorname{Var}[y_{\boldsymbol{x}} \mid \mathcal{D}_n]}{\operatorname{Var}[y_{\boldsymbol{x}} \mid f_{\boldsymbol{x}}, \mathcal{D}_n]} \right)
\end{aligned}
$$

$$
\begin{aligned}
&= 2\tilde{\sigma}^2 \cdot \mathrm{I}(f_{\boldsymbol{x}}; y_{\boldsymbol{x}} \mid \mathcal{D}_n) \\
&\overset{(iii)}{\leq} 2\tilde{\sigma}^2 \cdot \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}} \mid \mathcal{D}_n) \\
&\overset{(iv)}{\leq} 2\tilde{\sigma}^2 \cdot \max_{\boldsymbol{x}' \in \mathcal{S}} \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}'} \mid \mathcal{D}_n) \\
&\overset{(v)}{=} 2\tilde{\sigma}^2 \cdot \max_{\boldsymbol{x}' \in \mathcal{S}} \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}'} \mid \boldsymbol{y}_{1:n}) \\
&= 2\tilde{\sigma}^2 \cdot \Gamma_n
\end{aligned}
$$

where (i) follows from the noise assumption (cf. Assumption B.2); (ii) follows from Lemma C.38 and using monotonicity of variance; (iii) follows from $\boldsymbol{x} \in \mathcal{A}$ and monotonicity of information gain; (iv) follows from $\boldsymbol{x} \in \mathcal{S}$; and (v) uses that the posterior variance of Gaussians is independent of the realization and only depends on the *location* of observations. □

C.5.1 Convergence outside $\mathcal{S}$ for ITL

We will now show convergence of marginal variance to the irreducible uncertainty for points outside the sample space.

Our proof roughly proceeds as follows: We construct an "approximate Markov boundary" of $\boldsymbol{x}$ in $\mathcal{S}$, and show (1) that the size of this Markov boundary is independent of n, and (2) that a small uncertainty reduction within the Markov boundary implies that the marginal variances at the Markov boundary and(!) $\boldsymbol{x}$ are small.

Definition C.15 (Approximate Markov boundary)**.** For any $\epsilon > 0$, $n \geq 0$, and $\boldsymbol{x} \in \mathcal{X}$, we denote by $B_{n,\epsilon}(\boldsymbol{x})$ the smallest (multi-)subset of $\mathcal{S}$ such that

$$
\mathrm{Var}[f_{\boldsymbol{x}} \mid \mathcal{D}_n, \boldsymbol{y}_{B_{n,\epsilon}(\boldsymbol{x})}] \leq \eta_{\mathcal{S}}^2(\boldsymbol{x}) + \epsilon. \tag{13}
$$

We call $B_{n,\epsilon}(\boldsymbol{x})$ an *ϵ-approximate Markov boundary* of $\boldsymbol{x}$ in $\mathcal{S}$.

Equation (13) is akin to the notion of the smallest Markov blanket in $\mathcal{S}$ of some $\boldsymbol{x} \in \mathcal{X}$ (called a *Markov boundary*) which is the smallest set $\mathcal{B} \subseteq \mathcal{S}$ such that $f_{\boldsymbol{x}} \perp \boldsymbol{f}_{\mathcal{S}} \mid \boldsymbol{f}_{\mathcal{B}}$.

Lemma C.16 (Existence of an approximate Markov boundary)**.** *For any $\epsilon > 0$, let k be the smallest integer satisfying*

$$
\frac{\gamma_k}{k} \leq \frac{\epsilon \lambda_{\min}(\boldsymbol{K}_{\mathcal{S}\mathcal{S}})}{2|\mathcal{S}|\sigma^2\tilde{\sigma}^2}. \tag{14}
$$

Then, for any $n \geq 0$ and $\boldsymbol{x} \in \mathcal{X}$, there exists an ϵ-approximate Markov boundary $B_{n,\epsilon}(\boldsymbol{x})$ of $\boldsymbol{x}$ in $\mathcal{S}$ with size at most k.

Lemma C.16 shows that for any $\epsilon > 0$ there exists a universal constant b_ϵ such that

$$
|B_{n,\epsilon}(\boldsymbol{x})| \leq b_\epsilon \qquad \forall n \geq 0, \boldsymbol{x} \in \mathcal{X}. \tag{15}
$$

We defer the proof of Lemma C.16 to Appendix C.5.3 where we also provide an algorithm to compute $B_{n,\epsilon}(\boldsymbol{x})$.

Lemma C.17. *For any $\epsilon > 0$, $n \geq 0$, and $\boldsymbol{x} \in \mathcal{X}$,*

$$
\sigma_n^2(\boldsymbol{x}) \leq 2\sigma^2 \cdot \mathrm{I}(f_{\boldsymbol{x}}; \boldsymbol{y}_{B_{n,\epsilon}(\boldsymbol{x})} \mid \mathcal{D}_n) + \eta_{\mathcal{S}}^2(\boldsymbol{x}) + \epsilon \tag{16}
$$

where $B_{n,\epsilon}(\boldsymbol{x})$ is an ϵ-approximate Markov boundary of $\boldsymbol{x}$ in $\mathcal{S}$.

Proof. We have

$$
\begin{aligned}
\sigma_n^2(\boldsymbol{x}) &= \mathrm{Var}[f_{\boldsymbol{x}} \mid \mathcal{D}_n] - \eta_{\mathcal{S}}^2(\boldsymbol{x}) + \eta_{\mathcal{S}}^2(\boldsymbol{x}) \\
&\overset{(i)}{\leq} \mathrm{Var}[f_{\boldsymbol{x}} \mid \mathcal{D}_n] - \mathrm{Var}[f_{\boldsymbol{x}} \mid \boldsymbol{y}_{B_{n,\epsilon}(\boldsymbol{x})}, \mathcal{D}_n] + \eta_{\mathcal{S}}^2(\boldsymbol{x}) + \epsilon \\
&\overset{(ii)}{\leq} \sigma^2 \log\left(\frac{\mathrm{Var}[f_{\boldsymbol{x}} \mid \mathcal{D}_n]}{\mathrm{Var}[f_{\boldsymbol{x}} \mid \boldsymbol{y}_{B_{n,\epsilon}(\boldsymbol{x})}, \mathcal{D}_n]}\right) + \eta_{\mathcal{S}}^2(\boldsymbol{x}) + \epsilon
\end{aligned}
$$

$$= 2\sigma^2 \cdot \mathrm{I}(f_{\boldsymbol{x}}; \boldsymbol{y}_{B_{n,\epsilon}(\boldsymbol{x})} \mid \mathcal{D}_n) + \eta_{\mathcal{S}}^2(\boldsymbol{x}) + \epsilon$$

where (i) follows from the defining property of an ϵ-approximate Markov boundary (cf. Equation (13)); and (ii) follows from Lemma C.38 and using monotonicity of variance. □

Lemma C.18. *For any $\epsilon > 0, n \geq 0$, and $\boldsymbol{x} \in \mathcal{A}$,*

$$\mathrm{I}(f_{\boldsymbol{x}}; \boldsymbol{y}_{B_{n,\epsilon}(\boldsymbol{x})} \mid \mathcal{D}_n) \leq \frac{b_\epsilon}{\bar{\kappa}_n(B_{n,\epsilon}(\boldsymbol{x}))} \Gamma_n \tag{17}$$

where $B_{n,\epsilon}(\boldsymbol{x})$ is an ϵ-approximate Markov boundary of $\boldsymbol{x}$ in $\mathcal{S}$, $|B_{n,\epsilon}(\boldsymbol{x})| \leq b_\epsilon$, and where $\bar{\kappa}_n(\cdot) \stackrel{\text{def}}{=} \bar{\kappa}(\cdot \mid \boldsymbol{x}_{1:n})$ denotes the information ratio from Equation (7).

We remark that $\bar{\kappa}_n(\cdot) > 0$ as is shown in Lemma C.7, and hence, the right-hand side of the inequality is well-defined.

Proof. We use the abbreviated notation $B = B_{n,\epsilon}(\boldsymbol{x})$. We have

$$\begin{aligned}
\mathrm{I}(f_{\boldsymbol{x}}; \boldsymbol{y}_B \mid \mathcal{D}_n) &\stackrel{(i)}{\leq} \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; \boldsymbol{y}_B \mid \mathcal{D}_n) \\
&\stackrel{(ii)}{\leq} \frac{1}{\bar{\kappa}_{n,b_\epsilon}} \sum_{\tilde{\boldsymbol{x}} \in B} \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\tilde{\boldsymbol{x}}} \mid \mathcal{D}_n) \\
&\stackrel{(iii)}{\leq} \frac{b_\epsilon}{\bar{\kappa}_{n,b_\epsilon}} \max_{\tilde{\boldsymbol{x}} \in B} \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\tilde{\boldsymbol{x}}} \mid \mathcal{D}_n) \\
&\stackrel{(iv)}{\leq} \frac{b_\epsilon}{\bar{\kappa}_{n,b_\epsilon}} \max_{\tilde{\boldsymbol{x}} \in \mathcal{S}} \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\tilde{\boldsymbol{x}}} \mid \mathcal{D}_n) \\
&\stackrel{(v)}{=} \frac{b_\epsilon}{\bar{\kappa}_{n,b_\epsilon}} \max_{\tilde{\boldsymbol{x}} \in \mathcal{S}} \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\tilde{\boldsymbol{x}}} \mid \boldsymbol{y}_{1:n}) \\
&= \frac{b_\epsilon}{\bar{\kappa}_{n,b_\epsilon}} \Gamma_n
\end{aligned}$$

where (i) follows from monotonicity of mutual information; (ii) follows from the definition of the information ratio $\bar{\kappa}_{n,b_\epsilon}$ (cf. Equation (7)); (iii) follows from $b \leq b_\epsilon$; (iv) follows from $B \subseteq \mathcal{S}$; and (v) uses that the posterior variance of Gaussians is independent of the realization and only depends on the *location* of observations. □

Proof of Theorem 3.2 for ITL . The case where $\boldsymbol{x} \in \mathcal{A} \cap \mathcal{S}$ is shown by Lemma C.14 with $C = 2\tilde{\sigma}^2$.

To prove the more general result, fix any $\boldsymbol{x} \in \mathcal{A}$ and $\epsilon > 0$. By Lemma C.16, there exists an ϵ-approximate Markov boundary $B_{n,\epsilon}(\boldsymbol{x})$ of $\boldsymbol{x}$ in $\mathcal{S}$ such that $|B_{n,\epsilon}(\boldsymbol{x})| \leq b_\epsilon$. We have

$$\begin{aligned}
\sigma_n^2(\boldsymbol{x}) &\stackrel{(i)}{\leq} 2\sigma^2 \cdot \mathrm{I}(f_{\boldsymbol{x}}; \boldsymbol{y}_{B_{n,\epsilon}(\boldsymbol{x})} \mid \mathcal{D}_n) + \eta_{\mathcal{S}}^2(\boldsymbol{x}) + \epsilon \\
&\stackrel{(ii)}{\leq} \frac{2\sigma^2 b_\epsilon}{\bar{\kappa}_n(B_{n,\epsilon}(\boldsymbol{x}))} \Gamma_n + \eta_{\mathcal{S}}^2(\boldsymbol{x}) + \epsilon
\end{aligned}$$

where (i) follows from Lemma C.17; and (ii) follows from Lemma C.18.

Let $\epsilon = c\frac{\gamma_{\sqrt{n}}}{\sqrt{n}}$ with $c = 2|\mathcal{S}|\sigma^2\tilde{\sigma}^2/\lambda_{\min}(\boldsymbol{K}_{\mathcal{S}\mathcal{S}})$. Then, by Equation (14), b_ϵ can be bounded for instance by $\sqrt{n}$. Together with Theorem C.12 this implies for ITL that

$$\begin{aligned}
\nu_{n,\epsilon}^2 + \epsilon &\leq 2\sigma^2\sqrt{n}\,\Gamma_n + c\gamma_{\sqrt{n}}/\sqrt{n} \\
&\leq c'\gamma_n/\sqrt{n}
\end{aligned}$$

for a constant c', e.g., $c' = 2\sigma^2 + c$. □

C.5.2 Convergence outside $\mathcal{S}$ for VTL

Proof of Theorem 3.2 for VTL . Analogously to Lemma C.17, we have

$$\begin{aligned}\sigma_n^2(\boldsymbol{x}) &= \mathrm{Var}[f_{\boldsymbol{x}} \mid \mathcal{D}_n] - \eta_{\mathcal{S}}^2(\boldsymbol{x}) + \eta_{\mathcal{S}}^2(\boldsymbol{x}) \\ &\overset{(i)}{\leq} \mathrm{Var}[f_{\boldsymbol{x}} \mid \mathcal{D}_n] - \mathrm{Var}[f_{\boldsymbol{x}} \mid \boldsymbol{y}_{B_{n,\epsilon}(\boldsymbol{x})}, \mathcal{D}_n] + \eta_{\mathcal{S}}^2(\boldsymbol{x}) + \epsilon\end{aligned}$$

where (i) follows from the defining property of an ϵ-approximate Markov boundary (cf. Equation (13)). Further, we have

$$\begin{aligned}&\mathrm{Var}[f_{\boldsymbol{x}} \mid \mathcal{D}_n] - \mathrm{Var}[f_{\boldsymbol{x}} \mid \boldsymbol{y}_{B_{n,\epsilon}(\boldsymbol{x})}, \mathcal{D}_n] \\ &\overset{(i)}{\leq} \sum_{\tilde{\boldsymbol{x}} \in B_{n,\epsilon}(\boldsymbol{x})} (\mathrm{Var}[f_{\boldsymbol{x}} \mid \mathcal{D}_n] - \mathrm{Var}[f_{\boldsymbol{x}} \mid y_{\tilde{\boldsymbol{x}}}, \mathcal{D}_n]) \\ &\overset{(ii)}{\leq} \sum_{\tilde{\boldsymbol{x}} \in B_{n,\epsilon}(\boldsymbol{x})} (\mathrm{tr}\ \mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \boldsymbol{y}_{1:n}] - \mathrm{tr}\ \mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid y_{\tilde{\boldsymbol{x}}}, \boldsymbol{y}_{1:n}]) \\ &\overset{(iii)}{\leq} b_\epsilon \Gamma_n\end{aligned}$$

where (i) follows from the submodularity of $\psi_{\mathcal{A}}$; (ii) uses that the posterior variance of Gaussians is independent of the realization and only depends on the *location* of observations; and (iii) follows from the definition of Γ_n and Lemma C.16.

The remainder of the proof is analogous to the result for ITL, using Theorem C.13 to bound Γ_n. □

C.5.3 Existence of an Approximate Markov Boundary

We now derive Lemma C.16 which shows the existence of an approximate Markov boundary of $\boldsymbol{x}$ in $\mathcal{S}$.

Lemma C.19. *For any $S \subseteq \mathcal{S}$ and $k \geq 0$, there exists $B \subseteq S$ with $|B| = k$ such that for all $\boldsymbol{x'} \in S$,*

$$\mathrm{Var}[f_{\boldsymbol{x'}} \mid \boldsymbol{y}_B] \leq 2\tilde{\sigma}^2 \frac{\gamma_k}{k}. \tag{18}$$

Proof. We choose $B \subseteq S$ greedily using the acquisition function

$$\tilde{\boldsymbol{x}}_k \overset{\text{def}}{=} \arg\max_{\tilde{\boldsymbol{x}} \in S} \mathrm{I}(\boldsymbol{f}_S; y_{\tilde{\boldsymbol{x}}} \mid \boldsymbol{y}_{B_{k-1}})$$

where $B_k = \tilde{\boldsymbol{x}}_{1:k}$. Note that this is the "undirected" special case of ITL, and hence, we have

$$\begin{aligned}\mathrm{Var}\big[f_{\boldsymbol{x'}} \mid \boldsymbol{y}_{B_k}\big] &\overset{(i)}{\leq} 2\tilde{\sigma}^2 \Gamma_k \\ &\overset{(ii)}{\leq} 2\tilde{\sigma}^2 \frac{\gamma_k}{k}\end{aligned}$$

where (i) is due to Lemma C.14; and (ii) is due to Theorem C.12 and $\alpha_{S,S}(k) \leq 1$. □

Lemma C.20. *Given any $\epsilon > 0$ and $B \subseteq S \subseteq \mathcal{S}$ with $|S| < \infty$, such that for any $\boldsymbol{x'} \in S$,*

$$\mathrm{Var}[f_{\boldsymbol{x'}} \mid \boldsymbol{y}_B] \leq \frac{\epsilon \lambda_{\min}(\boldsymbol{K}_{SS})}{|S|\sigma^2}. \tag{19}$$

Then for any $\boldsymbol{x} \in \mathcal{X}$,

$$\mathrm{Var}[f_{\boldsymbol{x}} \mid \boldsymbol{y}_B] \leq \mathrm{Var}[f_{\boldsymbol{x}} \mid \boldsymbol{f}_S] + \epsilon. \tag{20}$$

Proof. We will denote the right-hand side of Equation (19) by ϵ'. We have

$$\begin{aligned}&\mathrm{Var}[f_{\boldsymbol{x}} \mid \boldsymbol{y}_B] \\ &\overset{(i)}{=} \mathbb{E}_{\boldsymbol{f}_S}[\mathrm{Var}_{f_{\boldsymbol{x}}}[f_{\boldsymbol{x}} \mid \boldsymbol{f}_S, \boldsymbol{y}_B] \mid \boldsymbol{y}_B] \\ &\qquad + \mathrm{Var}_{\boldsymbol{f}_S}[\mathbb{E}_{f_{\boldsymbol{x}}}[f_{\boldsymbol{x}} \mid \boldsymbol{f}_S, \boldsymbol{y}_B] \mid \boldsymbol{y}_B]\end{aligned}$$

$$\overset{(ii)}{=} \mathrm{Var}_{f_{\boldsymbol{x}}}[f_{\boldsymbol{x}} \mid \boldsymbol{f}_S, \boldsymbol{y}_B] + \mathrm{Var}_{\boldsymbol{f}_S}[\mathbb{E}_{f_{\boldsymbol{x}}}[f_{\boldsymbol{x}} \mid \boldsymbol{f}_S, \boldsymbol{y}_B] \mid \boldsymbol{y}_B]$$
$$\overset{(iii)}{=} \underbrace{\mathrm{Var}_{f_{\boldsymbol{x}}}[f_{\boldsymbol{x}} \mid \boldsymbol{f}_S]}_{\text{irreducible uncertainty}} + \underbrace{\mathrm{Var}_{\boldsymbol{f}_S}[\mathbb{E}_{f_{\boldsymbol{x}}}[f_{\boldsymbol{x}} \mid \boldsymbol{f}_S] \mid \boldsymbol{y}_B]}_{\text{reducible (epistemic) uncertainty}}$$

where (i) follows from the law of total variance; (ii) uses that the conditional variance of a Gaussian depends only on the location of observations and not on their value; and (iii) follows from $f_{\boldsymbol{x}} \perp \boldsymbol{y}_B \mid \boldsymbol{f}_S$ since $B \subseteq S$. It remains to bound the reducible uncertainty.

Let $h_{\boldsymbol{x}} : \mathbb{R}^d \to \mathbb{R},\ \boldsymbol{f}_S \mapsto \mathbb{E}[f_{\boldsymbol{x}} \mid \boldsymbol{f}_S]$ where we write $d \overset{\text{def}}{=} |S|$. Using the formula for the GP posterior mean, we have

$$h_{\boldsymbol{x}}(\boldsymbol{f}_S) = \mathbb{E}[f_{\boldsymbol{x}}] + \boldsymbol{z}^\top(\boldsymbol{f}_S - \mathbb{E}[\boldsymbol{f}_S])$$

where $\boldsymbol{z} \overset{\text{def}}{=} \boldsymbol{K}_{SS}^{-1}\boldsymbol{K}_{S\boldsymbol{x}}$. Because h is a linear function in $\boldsymbol{f}_S$ we have for the reducible uncertainty that

$$\begin{aligned}
\mathrm{Var}_{\boldsymbol{f}_S}[h_{\boldsymbol{x}}(\boldsymbol{f}_S) \mid \boldsymbol{y}_B] &= \boldsymbol{z}^\top \mathrm{Var}[\boldsymbol{f}_S \mid \boldsymbol{y}_B]\boldsymbol{z} \\
&\overset{(i)}{\leq} d \cdot \boldsymbol{z}^\top \mathrm{diag}\, \mathrm{Var}[\boldsymbol{f}_S \mid \boldsymbol{y}_B]\boldsymbol{z} \\
&\overset{(ii)}{\leq} \epsilon' d\, \boldsymbol{z}^\top \boldsymbol{z} \\
&= \epsilon' d\, \boldsymbol{K}_{\boldsymbol{x}S}\boldsymbol{K}_{SS}^{-1}\boldsymbol{K}_{SS}^{-1}\boldsymbol{K}_{S\boldsymbol{x}} \\
&\leq \frac{\epsilon' d}{\lambda_{\min}(\boldsymbol{K}_{SS})}\boldsymbol{K}_{\boldsymbol{x}S}\boldsymbol{K}_{SS}^{-1}\boldsymbol{K}_{S\boldsymbol{x}} \\
&\overset{(iii)}{\leq} \frac{\epsilon' d \sigma^2}{\lambda_{\min}(\boldsymbol{K}_{SS})}
\end{aligned}$$

where (i) follows from Lemma C.37; (ii) follows from the assumption that $\mathrm{Var}[f_{\boldsymbol{x}'} \mid \boldsymbol{y}_B] \leq \epsilon'$ for all $\boldsymbol{x}' \in S$; and (iii) follows from

$$\boldsymbol{K}_{\boldsymbol{x}S}\boldsymbol{K}_{SS}^{-1}\boldsymbol{K}_{S\boldsymbol{x}} \leq \boldsymbol{K}_{\boldsymbol{x}\boldsymbol{x}} = \sigma^2$$

since $\boldsymbol{K}_{\boldsymbol{x}\boldsymbol{x}} - \boldsymbol{K}_{\boldsymbol{x}S}\boldsymbol{K}_{SS}^{-1}\boldsymbol{K}_{S\boldsymbol{x}} \geq 0$. □

Proof of Lemma C.16. Let $B \subseteq \mathcal{S}$ be the set of size k generated by Lemma C.19 to satisfy $\mathrm{Var}[f_{\boldsymbol{x}'} \mid \boldsymbol{y}_B] \leq 2\tilde{\sigma}^2\gamma_k/k$ for all $\boldsymbol{x}' \in \mathcal{S}$. We have for any $\boldsymbol{x} \in \mathcal{X}$,

$$\begin{aligned}
\mathrm{Var}[f_{\boldsymbol{x}} \mid \mathcal{D}_n, \boldsymbol{y}_B] &\overset{(i)}{\leq} \mathrm{Var}[f_{\boldsymbol{x}} \mid \boldsymbol{y}_B] \\
&\overset{(ii)}{\leq} \mathrm{Var}[f_{\boldsymbol{x}} \mid \boldsymbol{f}_{\mathcal{S}}] + \epsilon
\end{aligned}$$

where (i) follows from monotonicity of variance; and (ii) follows from Lemma C.20; using $|\mathcal{S}| < \infty$ and the condition on k. □

We remark that Lemma C.19 provides an algorithm (just "undirected" ITL!) to compute an approximate Markov boundary, and the set B returned by this algorithm is a valid approximate Markov boundary for all $\boldsymbol{x} \in \mathcal{X}$. One can simply swap-in ITL with target space $\{\boldsymbol{x}\}$ for "undirected" ITL to obtain tighter (but instance-dependent) bounds on the size of the approximate Markov boundary.

C.5.4 Generalization to Continuous $\mathcal{S}$ for Finite Dimensional RKHSs

In this subsection we generalize Theorem 3.2 to continuous sample spaces $\mathcal{S}$. We will make the following assumption:

Assumption C.21. The RKHS of the kernel k is finite dimensional. In other words, the kernel k can be expressed as $k(\boldsymbol{x}, \boldsymbol{x}') = \boldsymbol{\phi}(\boldsymbol{x})^\top\boldsymbol{\phi}(\boldsymbol{x}')$ for some feature map $\boldsymbol{\phi} : \mathcal{X} \to \mathbb{R}^d$ with $d < \infty$.

In the following, we will denote the design matrix of the sample space $\mathcal{S}$ by $\boldsymbol{\Phi} \overset{\text{def}}{=} [\boldsymbol{\phi}(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{S}]^\top \in \mathbb{R}^{|\mathcal{S}| \times d}$, and we denote by $\boldsymbol{\Pi}_{\boldsymbol{\Phi}}$ its orthogonal projection onto the orthogonal complement of the span of $\boldsymbol{\Phi}$. In particular, it holds that

1. $\boldsymbol{\Pi}_{\boldsymbol{\Phi}} \boldsymbol{v} = \mathbf{0}$ for all $\boldsymbol{v} \in \operatorname{span} \boldsymbol{\Phi}$, and
2. $\boldsymbol{\Pi}_{\boldsymbol{\Phi}} \boldsymbol{v} = \boldsymbol{v}$ for all $\boldsymbol{v} \in (\operatorname{span} \boldsymbol{\Phi})^{\perp}$.

Especially, $\boldsymbol{v} \in \ker \boldsymbol{\Pi}_{\boldsymbol{\Phi}}$ if and only if $\boldsymbol{v} \in \operatorname{span} \boldsymbol{\Phi}$. This projection can be computed as follows:

Lemma C.22. *It holds that*

$$\boldsymbol{\Pi}_{\boldsymbol{\Phi}} = \boldsymbol{I} - \boldsymbol{\Phi}^{\top}(\boldsymbol{\Phi}\boldsymbol{\Phi}^{\top})^{-1}\boldsymbol{\Phi}. \tag{21}$$

Proof. $\boldsymbol{\Phi}^{\top}(\boldsymbol{\Phi}\boldsymbol{\Phi}^{\top})^{-1}\boldsymbol{\Phi}$ is the orthogonal projection onto the span of $\boldsymbol{\Phi}$ (see, e.g., **?**, page 211). □

Lemma C.23. *Under Assumption C.21, the irreducible uncertainty $\eta^2_{\mathcal{S}}(\boldsymbol{x})$ of $\boldsymbol{x} \in \mathcal{X}$ is*

$$\eta^2_{\mathcal{S}}(\boldsymbol{x}) = \|\boldsymbol{\phi}(\boldsymbol{x})\|^2_{\boldsymbol{\Pi}_{\boldsymbol{\Phi}}} \tag{22}$$

where $\|\boldsymbol{v}\|_{\boldsymbol{A}} = \sqrt{\boldsymbol{v}^{\top}\boldsymbol{A}\boldsymbol{v}}$ denotes the Mahalanobis distance.

Proof. This is an immediate consequence of the formula for the conditional variance of multivariate Gaussians (cf. Appendix B.2), applied to the linear kernel. □

Lemmas C.22 and C.23 imply that $\eta^2_{\mathcal{S}}(\boldsymbol{x}^{\parallel}) = 0$ for all $\boldsymbol{x}^{\parallel} \in \mathcal{X}$ with $\boldsymbol{\phi}(\boldsymbol{x}^{\parallel}) \in \operatorname{span} \boldsymbol{\Phi}$. That is, the irreducible uncertainty is zero for points in the span of the sample space. In contrast, for points $\boldsymbol{x}^{\perp}$ with $\boldsymbol{\phi}(\boldsymbol{x}^{\perp}) \in (\operatorname{span} \boldsymbol{\Phi})^{\perp}$, the irreducible uncertainty equals the initial uncertainty: $\eta^2_{\mathcal{S}}(\boldsymbol{x}^{\perp}) = \sigma^2_0(\boldsymbol{x}^{\perp})$. The irreducible uncertainty of any other point $\boldsymbol{x}$ can be computed by simple decomposition of $\boldsymbol{\phi}(\boldsymbol{x})$ into parallel and orthogonal components.

Assuming that Assumption C.21 holds and given any (non-finite) $\mathcal{S} \subseteq \mathcal{X}$, there exists a basis $\Omega_{\mathcal{S}} \subseteq \mathcal{X}$ in the space of embeddings $\boldsymbol{\phi}(\cdot)$ such that $\operatorname{span} \mathcal{S} = \operatorname{span} \Omega_{\mathcal{S}}$ and $|\Omega_{\mathcal{S}}| \leq d$. The generalized existence of an approximate Markov boundary for continuous domains can then be shown analogously to Lemma C.16:

Lemma C.24 (Existence of an approximate Markov boundary for a continuous domain)**.** *Let $\mathcal{S}$ be any (continuous) subset of $\mathcal{X}$ and let Assumption C.21 hold with $d < \infty$. Further, for any $\epsilon > 0$, let k be the smallest integer satisfying*

$$\frac{\gamma_k}{k} \leq \frac{\epsilon \lambda_{\min}(\boldsymbol{K}_{\Omega_{\mathcal{S}}\Omega_{\mathcal{S}}})}{2d\sigma^2\tilde{\sigma}^2}. \tag{23}$$

Then, for any $n \geq 0$ and $\boldsymbol{x} \in \mathcal{X}$, there exists an ϵ-approximate Markov boundary $B_{n,\epsilon}(\boldsymbol{x})$ of $\boldsymbol{x}$ in $\mathcal{S}$ with size at most k.

Proof sketch. The proof follows analogously to Lemma C.16 by conditioning on the finite set $\Omega_{\mathcal{S}}$ as opposed to $\mathcal{S}$. □

C.6 Proof of Theorem 3.3

We first formalize the assumptions of Theorem 3.3:

Assumption C.25 (Regularity of $f^{\star}$)**.** We assume that $f^{\star}$ is in a reproducing kernel Hilbert space $\mathcal{H}_k(\mathcal{X})$ associated with a kernel k and has bounded norm, that is, $\|f\|_k \leq B$ for some finite $B \in \mathbb{R}$.

Assumption C.26 (Sub-Gaussian noise)**.** We further assume that each ϵ_n from the noise sequence $\{\varepsilon_n\}_{n=1}^{\infty}$ is conditionally zero-mean $\rho(\boldsymbol{x}_n)$-sub-Gaussian with known constants $\rho(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \in \mathcal{X}$. Concretely,

$$\forall n \geq 1, \lambda \in \mathbb{R}: \quad \mathbb{E}\big[e^{\lambda \epsilon_n} \,\big|\, \mathcal{D}_{n-1}\big] \leq \exp\left(\frac{\lambda^2 \rho^2(\boldsymbol{x}_n)}{2}\right)$$

where $\mathcal{D}_{n-1}$ corresponds to the σ-algebra generated by the random variables $\{\boldsymbol{x}_i, \epsilon_i\}_{i=1}^{n-1}$ and $\boldsymbol{x}_n$.

We make use of the following foundational result, showing that under the above two assumptions the (misspecified) Gaussian process model from Section 3.1 is an all-time well-calibrated model of $f^{\star}$:

Lemma C.27 (Well-calibrated confidence intervals; Abbasi-Yadkori (2013); Chowdhury & Gopalan (2017))**.** *Pick $\delta \in (0,1)$ and let Assumptions C.25 and C.26 hold. Let*

$$\beta_n(\delta) = \|f^\star\|_k + \rho\sqrt{2(\gamma_n + 1 + \log(1/\delta))}$$

where $\rho = \max_{\boldsymbol{x}\in\mathcal{X}} \rho(\boldsymbol{x})$.[8] *Then, for all $\boldsymbol{x} \in \mathcal{X}$ and $n \geq 0$ jointly with probability at least $1-\delta$,*

$$|f^\star(\boldsymbol{x}) - \mu_n(\boldsymbol{x})| \leq \beta_n(\delta) \cdot \sigma_n(\boldsymbol{x})$$

where $\mu_n(\boldsymbol{x})$ and $\sigma_n^2(\boldsymbol{x})$ are mean and variance (as defined in Appendix B.2) of the GP posterior of $f(\boldsymbol{x})$ conditional on the observations $\mathcal{D}_n$, pretending that ε_i is Gaussian with variance $\rho^2(\boldsymbol{x}_i)$.

The proof of Theorem 3.3 is a straightforward application of Lemma C.27 and Theorem 3.2:

Proof of Theorem 3.3. By Theorem 3.2, we have that for all $\boldsymbol{x} \in \mathcal{A}$,

$$\sigma_n(\boldsymbol{x}) \leq \sqrt{\eta_{\mathcal{S}}^2(\boldsymbol{x}) + \nu_{\mathcal{A},\mathcal{S}}^2(n)} \leq \eta_{\mathcal{S}}(\boldsymbol{x}) + \nu_{\mathcal{A},\mathcal{S}}(n).$$

The result then follows by application of Lemma C.27. □

C.7 Proof of Theorem 4.1

In this section, we derive our main result on Safe BO. In Appendix C.7.1, we give the definition of the reachable safe set $\mathcal{R}$ and derive the conditions under which convergence to the reachable safe set is guaranteed. Then, in Appendix C.7.2, we prove Theorem 4.1.

Notation In the agnostic setting from Section 3.2 (i.e., under Assumptions C.25 and C.26), Lemma C.27 provides us with the following $(1-\delta)$-confidence intervals (CIs)

$$\mathcal{C}_n(\boldsymbol{x}) \stackrel{\text{def}}{=} \mathcal{C}_{n-1}(\boldsymbol{x}) \cap [\mu_n(\boldsymbol{x}) \pm \beta_n(\delta) \cdot \sigma_n(\boldsymbol{x})] \tag{24}$$

where $\mathcal{C}_{-1}(\boldsymbol{x}) = \mathbb{R}$. We write $u_n(\boldsymbol{x}) \stackrel{\text{def}}{=} \max \mathcal{C}_n(\boldsymbol{x})$, $l_n(\boldsymbol{x}) \stackrel{\text{def}}{=} \min \mathcal{C}_n(\boldsymbol{x})$, and $w_n(\boldsymbol{x}) \stackrel{\text{def}}{=} u_n(\boldsymbol{x}) - l_n(\boldsymbol{x})$ for its upper bound, lower bound, and width, respectively.

We learn separate statistical models f and $\{g_1, \dots, g_q\}$ for the ground truth objective $f^\star$ and ground truth constraints $\{g_1^\star, \dots, g_q^\star\}$. We write $\mathcal{I} \stackrel{\text{def}}{=} \{f, 1, \dots, q\}$ and collect the constraints in $\mathcal{I}_s \stackrel{\text{def}}{=} \{1, \dots, q\}$. Without loss of generality, we assume that the confidence intervals include the ground truths with probability at least $1-\delta$ jointly for all $i \in \mathcal{I}$.[9] For $i \in \mathcal{I}$, denote by $u_{n,i}, l_{n,i}, w_{n,i}, \eta_i, \beta_{n,i}$ the respective quantities. In the following, we do not explicitly denote the dependence of β_n on δ.

To improve clarity, we will refer to the set of potential maximizers defined in Equation (4) as $\mathcal{M}_n$ and denote by $\mathcal{A}_n$ an arbitrary target space.

We point out the following corollary:

Corollary C.28 (Safety)**.** *With high probability, jointly for any $n \geq 0$ and any $i \in \mathcal{I}_s$,*

$$\forall \boldsymbol{x} \in \mathcal{S}_n : g_i^\star(\boldsymbol{x}) \geq 0. \tag{25}$$

C.7.1 Convergence to Reachable Safe Set

Definition C.29 (Reachable safe set)**.** Given any pessimistic safe set $\mathcal{S} \subseteq \mathcal{X}$ and any $\epsilon \geq 0$ and $\beta \geq 0$, we define the *reachable safe set* up to (ϵ, β)-slack and its closure as

$$\begin{aligned}\mathcal{R}_{\epsilon,\beta}(\mathcal{S}) &\stackrel{\text{def}}{=} \mathcal{S} \cup \{\boldsymbol{x} \in \mathcal{X} \setminus \mathcal{S} \mid \\ &\qquad g_i^\star(\boldsymbol{x}) - \beta(\eta_i(\boldsymbol{x}; \mathcal{S}) + \epsilon) \geq 0 \text{ for all } i \in \mathcal{I}_s\} \\ \bar{\mathcal{R}}_{\epsilon,\beta}(\mathcal{S}) &\stackrel{\text{def}}{=} \lim_{n\to\infty} (\mathcal{R}_{\epsilon,\beta})^n(\mathcal{S})\end{aligned}$$

where $(\mathcal{R}_{\epsilon,\beta})^n$ denotes the n-th composition of $\mathcal{R}_{\epsilon,\beta}$ with itself.

[8] $\beta_n(\delta)$ can be tightened adaptively (Emmenegger et al., 2023).

[9] This can be achieved by taking a union bound and rescaling δ.

Remark C.30. Convergence of the safe set to the closure of the reachability operator can only be guaranteed for finite safe sets ($|\mathcal{S}^\star| < \infty$). The following proofs readily generalize to continuous domains by considering convergence within the k-th composition of the reachability operator with itself for some $k < \infty$. In this case the sample complexity grows with k rather than $|\mathcal{S}^\star|$. The only required modification is to lift the assumption of Theorem C.12 that information is gained only while safe sets remain constant (i.e., $\mathcal{S}_{i+1} = \mathcal{S}_i$ for all i). This assumption is straightforward to lift since for any $n \geq 0$ and $T \geq 1$,

$$\max_{\boldsymbol{x} \in \mathcal{S}_n} \Delta_{\mathcal{A}}(\boldsymbol{x} \mid \boldsymbol{x}_{1:n+T}) \leq \frac{1}{T} \sum_{t=1}^{T} \max_{\boldsymbol{x} \in \mathcal{S}_n} \Delta_{\mathcal{A}}(\boldsymbol{x} \mid \boldsymbol{x}_{1:n+t}) \leq \frac{1}{T} \sum_{t=1}^{T} \max_{\boldsymbol{x} \in \mathcal{S}_{n+t}} \Delta_{\mathcal{A}}(\boldsymbol{x} \mid \boldsymbol{x}_{1:n+t}) \leq \frac{\gamma_T}{T},$$

using submodularity for the first inequality and the monotonicity of the safe set for the second inequality. In particular, this shows that one continues learning about points in the original safe set — even as the safe set grows.

We denote by $\mathcal{S}_0$ the initial pessimistic safe set induced by the (prior) statistical model g (cf. Section 4) and write $\bar{\mathcal{R}}_{\epsilon,\beta} \overset{\text{def}}{=} \bar{\mathcal{R}}_{\epsilon,\beta}(\mathcal{S}_0)$.

Lemma C.31 (Properties of the reachable safe set)**.** *For all $\mathcal{S}, \mathcal{S}' \subseteq \mathcal{X}$, $\epsilon \geq 0$, and $\beta \geq 0$:*

(i) $\mathcal{S}' \subseteq \mathcal{S} \implies \mathcal{R}_{\epsilon,\beta}(\mathcal{S}') \subseteq \mathcal{R}_{\epsilon,\beta}(\mathcal{S})$,

(ii) $\mathcal{R}_{\epsilon,\beta}(\mathcal{S}) \subseteq \mathcal{S} \implies \bar{\mathcal{R}}_{\epsilon,\beta}(\mathcal{S}) \subseteq \mathcal{S}$, *and*

(iii) $\mathcal{R}_{0,0}(\emptyset) = \bar{\mathcal{R}}_{0,0} = \mathcal{S}^\star$.

Proof (adapted from lemma 7.1 of Berkenkamp et al. (2021)).

1. Let $\boldsymbol{x} \in \mathcal{R}_{\epsilon,\beta}(\mathcal{S}')$. If $\boldsymbol{x} \in \mathcal{S}$ then $\boldsymbol{x} \in \mathcal{R}_{\epsilon,\beta}(\mathcal{S})$, so let $\boldsymbol{x} \notin \mathcal{S}$. Then, by definition, for all $i \in \mathcal{I}_s$, $f_i^\star(\boldsymbol{x}) - \beta\eta_i(\boldsymbol{x}; \mathcal{S}') - \epsilon \geq 0$. By the monotonicity of variance, $\eta_i(\boldsymbol{x}; \mathcal{S}') \geq \eta_i(\boldsymbol{x}; \mathcal{S})$ for all $i \in \mathcal{I}$, and hence $f_i^\star(\boldsymbol{x}) - \beta\eta_i(\boldsymbol{x}; \mathcal{S}) - \epsilon \geq 0$ for all $i \in \mathcal{I}_s$. It follows that $\boldsymbol{x} \in \mathcal{R}_{\epsilon,\beta}(\mathcal{S})$.

2. By the monotonicity of variance, $\eta_i(\boldsymbol{x}; \mathcal{R}_{\epsilon,\beta}(\mathcal{S})) \geq \eta_i(\boldsymbol{x}; \mathcal{S})$ for all $\boldsymbol{x} \in \mathcal{X}$ and $i \in \mathcal{I}$. Thus, by definition of the safe region, we have that $\mathcal{R}_{\epsilon,\beta}(\mathcal{R}_{\epsilon,\beta}(\mathcal{S})) \subseteq \mathcal{S}$. The result follows by taking the limit.

3. The result follows directly from the definition of the true safe set $\mathcal{S}^\star$ (cf. Equation (3)). □

Clearly, we cannot expand the safe set beyond $\bar{\mathcal{R}}_{0,0}$. The following is our main intermediate result, showing that either we expand the safe set at some point or the uncertainty converges to the irreducible uncertainty.

Lemma C.32. *Given any $n_0 \geq 0$, $\epsilon > 0$, let n' be the smallest integer such that $\nu_{n',\tilde{\epsilon}^2} \leq \tilde{\epsilon}$ where $\tilde{\epsilon} = \epsilon/2$. Let $\beta_{n_0+n'} = \max_{i \in \mathcal{I}_s} \beta_{n_0+n',i}$. Assume that the sequence of target spaces is monotonically decreasing, i.e., $\mathcal{A}_{n+1} \subseteq \mathcal{A}_n$. Then, we have with high probability (at least) one of*

$$\Big(\forall \boldsymbol{x} \in \mathcal{A}_{n_0+n'},\ \forall i \in \mathcal{I}: \\ w_{n_0+n',i}(\boldsymbol{x}) \leq \beta_{n_0+n'}[\eta_i(\boldsymbol{x}; \mathcal{S}_{n_0+n'}) + \epsilon] \\ \textit{and} \quad \mathcal{A}_{n_0+n'} \cap \mathcal{R}_{\epsilon,\beta_{n_0+n'}}(\mathcal{S}_{n_0+n'}) \subseteq \mathcal{S}_{n_0+n'}\Big)$$

or $|\mathcal{S}_{n_0+n'+1}| > |\mathcal{S}_{n_0}|$.

Proof. Suppose that $|\mathcal{S}_{n_0+n'+1}| = |\mathcal{S}_{n_0}|$. Then, by Theorem 3.3 (using that the sequence of target spaces is monotonically decreasing), for any $\boldsymbol{x} \in \mathcal{A}_{n_0+n'}$ and $i \in \mathcal{I}$,

$$w_{n_0+n',i}(\boldsymbol{x}) \leq \beta_{n_0+n'}[\eta_i(\boldsymbol{x}; \mathcal{S}_{n_0+n'}) + \epsilon].$$

As $\mathcal{S}_{n_0+n'+1} = \mathcal{S}_{n_0+n'}$ we have for all $\boldsymbol{x} \in \mathcal{A}_{n_0+n'} \setminus \mathcal{S}_{n_0+n'}$ and $i \in \mathcal{I}_s$, with high probability that

$$\begin{aligned} 0 > l_{n_0+n',i}(\boldsymbol{x}) &\geq g_i^\star(\boldsymbol{x}) - w_{n_0+n',i}(\boldsymbol{x}) \\ &\geq g_i^\star(\boldsymbol{x}) - \beta_{n_0+n'}[\eta_i(\boldsymbol{x}; \mathcal{S}_{n_0+n'}) + \epsilon]. \end{aligned}$$

It follows that $\mathcal{A}_{n_0+n'} \cap \mathcal{R}_{\epsilon,\beta_{n_0+n'}}(\mathcal{S}_{n_0+n'}) \subseteq \mathcal{S}_{n_0+n'}$. □

To gather more intuition about the above lemma, consider the target space

$$\mathcal{E}_n \stackrel{\text{def}}{=} \widehat{\mathcal{S}}_n \setminus \mathcal{S}_n. \tag{26}$$

We call $\mathcal{E}_n$ the *potential expanders* since it contains all points which might be safe, but are not yet known to be safe. Under this target space, the above lemma simplifies slightly:

Lemma C.33. *For any $n \geq 0$ and $\epsilon, \beta \geq 0$, if $\mathcal{E}_n \subseteq \mathcal{A}_n$ then with high probability,*

$$\mathcal{S}_n \cup (\mathcal{A}_n \cap \mathcal{R}_{\epsilon,\beta}(\mathcal{S}_n)) = \mathcal{R}_{\epsilon,\beta}(\mathcal{S}_n).$$

Proof. With high probability, $\mathcal{R}_{\epsilon,\beta}(\mathcal{S}_n) \subseteq \widehat{\mathcal{S}}_n = \mathcal{S}_n \cup \mathcal{E}_n$. The lemma is a direct consequence. □

The above lemmas can be combined to yield our main result of this subsection, establishing the convergence of ITL to the reachable safe set.

Theorem C.34 (Convergence to reachable safe set)**.** *For any $\epsilon > 0$, let n' be the smallest integer satisfying the condition of Lemma C.32, and define $n^\star \stackrel{\text{def}}{=} (|\mathcal{S}^\star| + 1)n'$. Let $\bar{\beta}_{n^\star} \geq \beta_{n,i}$ for all $n \leq n^\star, i \in \mathcal{I}_s$. Assume that the sequence of target spaces is monotonically decreasing, i.e., $\mathcal{A}_{n+1} \subseteq \mathcal{A}_n$. Then, the following inequalities hold jointly with probability at least $1 - \delta$:*

(i) $\forall n \geq 0,\ \forall i \in \mathcal{I}_s : g_i^\star(\boldsymbol{x}_n) \geq 0,$

safety

(ii) $\mathcal{A}_{n^\star} \cap \bar{\mathcal{R}}_{\epsilon,\bar{\beta}_{n^\star}} \subseteq \mathcal{S}_{n^\star} \subseteq \bar{\mathcal{R}}_{0,0} = \mathcal{S}^\star,$

convergence to safe region

(iii) $\forall \boldsymbol{x} \in \mathcal{A}_{n^\star},\ \forall i \in \mathcal{I} : w_{n^\star,i}(\boldsymbol{x}) \leq \bar{\beta}_{n^\star}\eta_i(\boldsymbol{x}; \bar{\mathcal{R}}_{\epsilon,\bar{\beta}_{n^\star}}) + \epsilon,$

convergence of width

(iv) $\forall \boldsymbol{x} \in \bar{\mathcal{R}}_{\epsilon,\bar{\beta}_{n^\star}},\ \forall i \in \mathcal{I} : \eta_i(\boldsymbol{x}; \bar{\mathcal{R}}_{\epsilon,\bar{\beta}_{n^\star}}) = 0.$

convergence of width within safe region

Proof. (i) is a direct consequence of Corollary C.28. $\mathcal{S}_{n^\star} \subseteq \mathcal{S}^\star$ follows directly from the pessimistic safe set $\mathcal{S}_{n^\star}$ from (ii) being a subset of the true safe set $\mathcal{S}^\star$. (iv) follows directly from the definition of irreducible uncertainty. Thus, it remains to establish $\mathcal{A}_{n^\star} \cap \bar{\mathcal{R}}_{\epsilon,\bar{\beta}_{n^\star}} \subseteq \mathcal{S}_{n^\star}$ and (iii).

Recall that with high probability $|\mathcal{S}_n| \in [0, |\mathcal{S}^\star|]$ for all $n \geq 0$. Thus, the size of the pessimistic safe set can increase at most $|\mathcal{S}^\star|$ many times. By Lemma C.32, using the assumption on n', the size of the pessimistic safe set increases at least once every n' iterations, or else:

$$\begin{aligned} &\forall \boldsymbol{x} \in \mathcal{A}_{n_0+n'},\ \forall i \in \mathcal{I} : w_{n_0+n',i}(\boldsymbol{x}) \leq \beta_{n_0+n'}[\eta_i(\boldsymbol{x}; \mathcal{S}_{n_0+n'}) + \epsilon] \\ &\text{and} \quad \mathcal{A}_{n_0+n'} \cap \mathcal{R}_{\epsilon,\beta_{n_0+n'}}(\mathcal{S}_{n_0+n'}) \subseteq \mathcal{S}_{n_0+n'}. \end{aligned} \tag{27}$$

Because the safe set can expand at most $|\mathcal{S}^\star|$ many times, Equation (27) occurs eventually for some $n_0 \leq |\mathcal{S}^\star|n'$. In this case, since $\bar{\beta}_{n^\star} \geq \beta_{n_0+n'}$ and $\mathcal{A}_{n^\star} \subseteq \mathcal{A}_{n_0+n'}$ (as $n_0 + n' \leq n^\star$) we have that

$$\begin{aligned} \mathcal{A}_{n^\star} \cap \mathcal{R}_{\epsilon,\bar{\beta}_{n^\star}}(\mathcal{S}_{n_0+n'}) &\subseteq \mathcal{A}_{n_0+n'} \cap \mathcal{R}_{\epsilon,\beta_{n_0+n'}}(\mathcal{S}_{n_0+n'}) \\ &\subseteq \mathcal{S}_{n_0+n'}. \end{aligned}$$

By Lemma C.31 (ii), this implies

$$\mathcal{A}_{n^\star} \cap \bar{\mathcal{R}}_{\epsilon,\bar{\beta}_{n^\star}} \subseteq \mathcal{S}_{n_0+n'} \subseteq \mathcal{S}_{n^\star}.$$

□

We emphasize that Theorem C.34 holds for arbitrary target spaces $\mathcal{A}_n$. If additionally, $\mathcal{E}_n \subseteq \mathcal{A}_n$ for all $n \geq 0$ then by Lemma C.33, Theorem C.34 (ii) strengthens to $\bar{\mathcal{R}}_{\epsilon,\bar{\beta}_{n^\star}} \subseteq \mathcal{S}_{n^\star}$. Intuitively, $\mathcal{E}_n \subseteq \mathcal{A}_n$ ensures that one aims to expand the safe set in *all* directions. Conversely, if $\mathcal{E}_n \not\subseteq \mathcal{A}_n$ then one aims only to expand the safe set in the direction of $\mathcal{A}_n$ (or not at all if $\mathcal{A}_n \subseteq \mathcal{S}_n$).

"Free" convergence guarantees in many applications Theorem C.34 can be specialized to yield convergence guarantees in various settings by choosing an appropriate target space $\mathcal{A}_n$. Straightforward application of Theorem C.34 (informally) requires that the sequence of target spaces is monotonically decreasing (i.e., $\mathcal{A}_{n+1} \subseteq \mathcal{A}_n$), and that each target space $\mathcal{A}_n$ is an "over-approximation" of the actual set of targeted points (such as the set of optimas in the Bayesian optimization setting). We discuss two such applications in the following.

1. *Pure expansion:* For example, for the target space $\mathcal{E}_n$, Theorem C.34 bounds the convergence of the safe set to the reachable safe set. In this case, the transductive active learning problem corresponds to the "pure expansion" setting, also addressed by the ISE baseline discussed in Section 4. The ISE baseline, however, does not establish convergence guarantees of the kind of Theorem C.34. Note that $\mathcal{E}_n$ satisfies the (informal) requirements laid out previously, since it is monotonically decreasing by definition, and with high probability, any point $\boldsymbol{x} \in \mathcal{S}^\star$ that is not in $\mathcal{S}_n$ is contained within $\mathcal{E}_n$.
2. *Level set estimation:* Given any $\tau \in \mathbb{R}$, we denote the (safe) τ-level set of $f^\star$ by $\mathcal{L}^\tau \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \mathcal{S}^\star \mid f^\star(\boldsymbol{x}) = \tau\}$. We define the *potential level set* as
$$\mathcal{L}_n^\tau \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \widehat{\mathcal{S}}_n \mid l_n^f(\boldsymbol{x}) \leq \tau \leq u_n^f(\boldsymbol{x})\}. \tag{28}$$
That is, $\mathcal{L}_n^\tau$ is the subset of the optimistic safe set $\widehat{\mathcal{S}}_n$ where the τ-level set of $f^\star$ may be located. Analogously to the potential expanders, it is straightforward to show that $\mathcal{L}_n^\tau$ over-approximates the true τ-level set and is monotonically decreasing.

We remark that our guarantees from this section also apply to the standard ("unsafe") setting where $\mathcal{S}^\star = \mathcal{S}_0 = \mathcal{X}$.

C.7.2 Convergence to Safe Optimum

In this section, we specialize Theorem C.34 for the case that the target space contains the potential maximizers $\mathcal{M}_n$ (cf. Equation (4)). It is straightforward to see that the sequence $\mathcal{M}_n$ is monotonically decreasing (i.e., $\mathcal{M}_{n+1} \subseteq \mathcal{M}_n$). The following lemma shows that the potential maximizers over-approximate the set of safe maxima $\mathcal{X}^* \stackrel{\text{def}}{=} \arg\max_{\boldsymbol{x} \in \mathcal{S}^\star} f^\star(\boldsymbol{x})$.

Lemma C.35 (Potential maximizers over-approximate safe maxima)**.** *For all $n \geq 0$ and with probability at least $1 - \delta$,*

(i) $\boldsymbol{x} \in \mathcal{X}^*$ implies $\boldsymbol{x} \in \mathcal{M}_n$ and

(ii) $\boldsymbol{x} \notin \mathcal{M}_n$ implies $\boldsymbol{x} \notin \mathcal{X}^*$.

Proof. If $\boldsymbol{x} \notin \mathcal{M}_n$ then
$$u_{n,f}(\boldsymbol{x}) < \max_{\boldsymbol{x}' \in \mathcal{S}_n} l_{n,f}(\boldsymbol{x}') \leq \max_{\boldsymbol{x}' \in \mathcal{S}^\star} l_{n,f}(\boldsymbol{x}')$$
where we used $\mathcal{S}_n \subseteq \mathcal{S}^\star$ with high probability, which directly implies with high probability that $\boldsymbol{x} \notin \mathcal{X}^*$.

For the other direction, if $\boldsymbol{x} \in \mathcal{X}^*$ then
$$u_{n,f}(\boldsymbol{x}) \geq \max_{\boldsymbol{x}' \in \mathcal{S}^\star} l_{n,f}(\boldsymbol{x}') \geq \max_{\boldsymbol{x}' \in \mathcal{S}_n} l_{n,f}(\boldsymbol{x}')$$
with high probability. □

We denote the set of optimal actions which are safe up to (ϵ, β)-slack by
$$\mathcal{X}^*_{\epsilon,\beta} \stackrel{\text{def}}{=} \arg\max_{\boldsymbol{x} \in \bar{\mathcal{R}}_{\epsilon,\beta}} f^\star(\boldsymbol{x}),$$
and by $f^*_{\epsilon,\beta}$ the maximum value attained by $f^\star$ at any of the points in $\mathcal{X}^*_{\epsilon,\beta}$. The regret can be expressed as
$$r_n(\bar{\mathcal{R}}_{\epsilon,\beta}) = f^*_{\epsilon,\beta} - f^\star(\widehat{\boldsymbol{x}}_n)$$
The following theorem formalizes Theorem 4.1 and establishes convergence to the safe optimum.

Theorem C.36 (Convergence to safe optimum)**.** *For any $\epsilon > 0$, let n' be the smallest integer satisfying the condition of Lemma C.32, and define $n^\star \stackrel{\text{def}}{=} (|\mathcal{S}^\star| + 1)n'$. Let $\bar{\beta}_{n^\star} \geq \beta_{n,i}$ for all $n \leq n^\star, i \in \mathcal{I}_s$. Then, the following inequalities hold jointly with probability at least $1 - \delta$:*

(i) $\forall n \geq 0,\ \forall i \in \mathcal{I}_s : g_i^\star(\boldsymbol{x}_n) \geq 0,$

safety

(ii) $\forall n \geq n^\star : r_n(\bar{\mathcal{R}}_{\epsilon,\bar{\beta}_{n^\star}}) \leq \epsilon.$

convergence to safe optimum

Proof. Fix any $\boldsymbol{x}^* \in \mathcal{X}^*_{\epsilon,\bar{\beta}_{n^\star}} \subseteq \bar{\mathcal{R}}_{\epsilon,\bar{\beta}_{n^\star}}$. Assume w.l.o.g. that $\boldsymbol{x}^* \in \mathcal{M}_{n^\star}$.[10] Then, with high probability,

$$\begin{aligned} f^*_{\epsilon,\bar{\beta}_{n^\star}} = f^\star(\boldsymbol{x}^*) &\leq u_{n^\star,f}(\boldsymbol{x}^*) \\ &= l_{n^\star,f}(\boldsymbol{x}^*) + w_{n^\star,f}(\boldsymbol{x}^*) \\ &\stackrel{(i)}{\leq} l_{n^\star,f}(\widehat{\boldsymbol{x}}_{n^\star}) + w_{n^\star,f}(\boldsymbol{x}^*) \\ &\leq f^\star(\widehat{\boldsymbol{x}}_{n^\star}) + w_{n^\star,f}(\boldsymbol{x}^*) \\ &\stackrel{(ii)}{\leq} f^\star(\widehat{\boldsymbol{x}}_{n^\star}) + \epsilon \end{aligned}$$

where (i) follows from the definition of $\widehat{\boldsymbol{x}}_n$; and (ii) follows from Theorem C.34 and noting that $\boldsymbol{x}^* \in \mathcal{M}_{n^\star} \cap \bar{\mathcal{R}}_{\epsilon,\bar{\beta}_{n^\star}}$.

We have shown that $f^\star(\widehat{\boldsymbol{x}}_{n^\star}) \geq f^*_{\epsilon,\bar{\beta}_{n^\star}} - \epsilon$, which implies $r_{n^\star}(\bar{\mathcal{R}}_{\epsilon,\bar{\beta}_{n^\star}}) \leq \epsilon$. Since the upper- and lower-confidence bounds are monotonically decreasing / increasing, respectively, we have that for all $n \geq n^\star$, $r_n(\bar{\mathcal{R}}_{\epsilon,\bar{\beta}_{n^\star}}) \leq \epsilon$. □

C.8 Useful Facts and Inequalities

We denote by $\preceq$ the Loewner partial ordering of symmetric matrices.

Lemma C.37. *Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ be a positive definite matrix with diagonal $\boldsymbol{D}$. Then, $\boldsymbol{A} \preceq n\boldsymbol{D}$.*

Proof. Equivalently, one can show $n\boldsymbol{D} - \boldsymbol{A} \succeq \boldsymbol{0}$. We write $\boldsymbol{A} \stackrel{\text{def}}{=} \boldsymbol{D}^{1/2}\boldsymbol{Q}\boldsymbol{D}^{1/2}$, and thus, $\boldsymbol{Q} = \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}$ is a positive definite symmetric matrix with all diagonal elements equal to 1. It remains to show that

$$n\boldsymbol{D} - \boldsymbol{A} = \boldsymbol{D}^{1/2}(n\boldsymbol{I} - \boldsymbol{Q})\boldsymbol{D}^{1/2} \succeq \boldsymbol{0}.$$

Note that $\sum_{i=1}^n \lambda_i(\boldsymbol{Q}) = \operatorname{tr} \boldsymbol{Q} = n$, and hence, all eigenvalues of $\boldsymbol{Q}$ belong to $(0, n)$. □

Lemma C.38. *If $a, b \in (0, M]$ for some $M > 0$ and $b \geq a$ then*

$$b - a \leq M \cdot \log\left(\frac{b}{a}\right). \tag{29}$$

If additionally, $a \geq M'$ for some $M' > 0$ then

$$b - a \geq M' \cdot \log\left(\frac{b}{a}\right). \tag{30}$$

Proof. Let $f(x) \stackrel{\text{def}}{=} \log x$. By the mean value theorem, there exists $c \in (a, b)$ such that

$$\frac{1}{c} = f'(c) = \frac{f(b) - f(a)}{b - a} = \frac{\log b - \log a}{b - a} = \frac{\log(\frac{b}{a})}{b - a}.$$

Thus,

$$b - a = c \cdot \log\left(\frac{b}{a}\right) < M \cdot \log\left(\frac{b}{a}\right).$$

[10] Otherwise, with high probability, $f^\star(\widehat{\boldsymbol{x}}_{n^\star}) > f^*_{\epsilon,\bar{\beta}_{n^\star}}$.

Under the additional condition that $a \geq M'$, we obtain

$$b - a = c \cdot \log\left(\frac{b}{a}\right) > M' \cdot \log\left(\frac{b}{a}\right).$$

□

D Interpretations & Approximations of Principle (†)

We give a brief overview of interpretations and approximations of ITL, as well as alternative decision rules adhering to the fundamental principle (†).

The discussed interpretations of (†) differ mainly in how they quantify the "uncertainty" about $\mathcal{A}$. In the GP setting, this "uncertainty" is captured by the covariance matrix $\boldsymbol{\Sigma}$ of $\boldsymbol{f}_{\mathcal{A}}$, and we consider two main ways of "scalarizing" $\boldsymbol{\Sigma}$:

1. the total (marginal) variance $\operatorname{tr} \boldsymbol{\Sigma}$, and
2. the "generalized variance" $|\boldsymbol{\Sigma}|$.

The generalized variance — which was originally suggested by Wilks (1932) as a generalization of variance to multiple dimensions — takes into account correlations. In contrast, the total variance discards all correlations between points in $\mathcal{A}$.

All discussed decision rules following principle (†) (i.e., ITL, VTL, MM-ITL) differ only in their weighting of the points in $\mathcal{A}$, and they coincide when $|\mathcal{A}| = 1$.

D.1 Interpretations of ITL

We briefly discuss three interpretations of ITL.

Minimizing generalized variance In the GP setting, ITL can be equivalently characterized as minimizing generalized posterior variance:

$$\begin{aligned} \boldsymbol{x}_n &= \arg\max_{\boldsymbol{x} \in \mathcal{S}} \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}} \mid \mathcal{D}_n) \\ &= \arg\max_{\boldsymbol{x} \in \mathcal{S}} \frac{1}{2} \log\left(\frac{|\mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_{n-1}]|}{|\mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_{n-1}, y_{\boldsymbol{x}}]|}\right) \\ &= \arg\min_{\boldsymbol{x} \in \mathcal{S}} |\mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_{n-1}, y_{\boldsymbol{x}}]|\,. \end{aligned} \tag{31}$$

Maximizing relevance and minimizing redundancy An alternative interpretation of ITL is

$$\mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}} \mid \mathcal{D}_n) = \underbrace{\mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}})}_{\text{relevance}} - \underbrace{\mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}}; \mathcal{D}_n)}_{\text{redundancy}} \tag{32}$$

where $\mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}}; \mathcal{D}_n) = \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}}) - \mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}} \mid \mathcal{D}_n)$ denotes the *multivariate information gain* (cf. Appendix B). In this way, ITL can be seen as maximizing observation relevance while minimizing observation redundancy. This interpretation is common in the literature on feature selection (Peng et al., 2005; Vergara & Estévez, 2014; Beraha et al., 2019).

Steepest descent in measure spaces ITL can be seen as performing steepest descent in the space of probability measures over $\boldsymbol{f}_{\mathcal{A}}$, with the KL divergence as metric:

$$\mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}} \mid \mathcal{D}_n) = \mathbb{E}_{y_{\boldsymbol{x}}}[\mathrm{KL}(p(\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_n, y_{\boldsymbol{x}}) \| p(\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_n))].$$

That is, ITL finds the observation yielding the "largest update" to the current density.

D.2 Interpretations of VTL

Quantifying the uncertainty about $\boldsymbol{f}_{\mathcal{A}}$ by the marginal variance of points in $\mathcal{A}$ rather than entropy (or generalized variance), the principle (†) leads to VTL. Note that if $|\mathcal{A}| = 1$, then VTL is equivalent to ITL. Unlike the similar, but more technical, TruVar algorithm proposed by Bogunovic et al. (2016), VTL does not require truncated variances, and hence, VTL can be applied to constrained settings (where $\mathcal{A} \not\subseteq \mathcal{S}$) as well.

Relationship to ITL Note that the ITL criterion in the GP setting can be expressed as

$$\boldsymbol{x}_n = \arg\min_{\boldsymbol{x}\in\mathcal{S}} \operatorname{tr} \log \operatorname{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_{n-1}, y_{\boldsymbol{x}}] \tag{33}$$

where for a positive semi-definite matrix $\boldsymbol{A}$ with spectral decomposition $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^\top$ we write $\log \boldsymbol{A} = \boldsymbol{V} \log \boldsymbol{\Lambda} \boldsymbol{V}^\top$ for the logarithmic matrix function. To derive Equation (33) we use that $\log |\boldsymbol{A}| = \sum_i \log \lambda_i(\boldsymbol{A}) = \operatorname{tr} \log \boldsymbol{A}$. Hence, ITL and VTL are identical up to a different weighting of the eigenvalues of the posterior covariance matrix.

Minimizing a bound to the approximation error Chowdhury & Gopalan (2017) (page 19) bound the approximation error $|f^\star(\boldsymbol{x}) - \mu_n(\boldsymbol{x})|$ by

$$\underbrace{|\boldsymbol{k}_t(\boldsymbol{x})^\top (\boldsymbol{K}_t + \boldsymbol{P}_t)^{-1} \boldsymbol{\varepsilon}_{1:t}|}_{\text{variance}} + \underbrace{|f^\star(\boldsymbol{x}) - \boldsymbol{k}_t(\boldsymbol{x})^\top (\boldsymbol{K}_t + \boldsymbol{P}_t)^{-1} \boldsymbol{f}_{1:t}|}_{\text{bias}}$$

where $\boldsymbol{k}_t(\boldsymbol{x}) \stackrel{\text{def}}{=} \boldsymbol{K}_{\boldsymbol{x}\boldsymbol{x}_{1:t}}$, $\boldsymbol{K}_t \stackrel{\text{def}}{=} \boldsymbol{K}_{\boldsymbol{x}_{1:t}\boldsymbol{x}_{1:t}}$, and $\boldsymbol{P}_t \stackrel{\text{def}}{=} \boldsymbol{P}_{\boldsymbol{x}_{1:t}}$. Similar to a standard bias-variance decomposition, the first term measures variance and the second term measures bias. Following Lemma C.27, VTL can be seen as greedily minimizing this bound to the approximation error (i.e., both bias and variance).

Maximizing correlation to prediction targets weighted by their variance It can be shown (see the proof below) that the VTL decision rule is equivalent to

$$\boldsymbol{x}_n = \arg\max_{\boldsymbol{x}\in\mathcal{S}} \sum_{\boldsymbol{x}'\in\mathcal{A}} \operatorname{Var}[f_{\boldsymbol{x}'} \mid \mathcal{D}_{n-1}] \cdot \operatorname{Cor}[f_{\boldsymbol{x}'}, y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1}]^2. \tag{34}$$

That is, VTL maximizes the squared correlation between the next observation and the prediction targets, weighted by their variance. Intuitively, prediction targets are weighted by their variance since more can be learned about a prediction target with higher variance. This is precisely what leads to the "diverse" sample selection, and is akin to "uncertainty sampling" among the prediction targets and then selecting the observation which is most correlated with the selected prediction target.

Proof. Starting with the VTL objective, we have

$$\begin{aligned}
\arg\min_{\boldsymbol{x}\in\mathcal{S}} \sum_{\boldsymbol{x}'\in\mathcal{A}} \operatorname{Var}[f_{\boldsymbol{x}'} \mid \mathcal{D}_n, y_{\boldsymbol{x}}] &= \arg\min_{\boldsymbol{x}\in\mathcal{S}} \sum_{\boldsymbol{x}'\in\mathcal{A}} \left(\operatorname{Var}[f_{\boldsymbol{x}'} \mid \mathcal{D}_n] - \frac{\operatorname{Cov}[f_{\boldsymbol{x}'}, y_{\boldsymbol{x}} \mid \mathcal{D}_n]^2}{\operatorname{Var}[y_{\boldsymbol{x}} \mid \mathcal{D}_n]} \right) \\
&= \arg\max_{\boldsymbol{x}\in\mathcal{S}} \sum_{\boldsymbol{x}'\in\mathcal{A}} \frac{\operatorname{Var}[f_{\boldsymbol{x}'} \mid \mathcal{D}_n] \cdot \operatorname{Cov}[f_{\boldsymbol{x}'}, y_{\boldsymbol{x}} \mid \mathcal{D}_n]^2}{\operatorname{Var}[f_{\boldsymbol{x}'} \mid \mathcal{D}_n] \cdot \operatorname{Var}[y_{\boldsymbol{x}} \mid \mathcal{D}_n]} + \text{const} \\
&= \arg\max_{\boldsymbol{x}\in\mathcal{S}} \sum_{\boldsymbol{x}'\in\mathcal{A}} \operatorname{Var}[f_{\boldsymbol{x}'} \mid \mathcal{D}_n] \cdot \operatorname{Cor}[f_{\boldsymbol{x}'}, y_{\boldsymbol{x}} \mid \mathcal{D}_n]^2 + \text{const}.
\end{aligned}$$

□

D.3 Mean Marginal ITL

MacKay (1992) previously proposed "mean-marginal" ITL (MM-ITL) in the setting where $\mathcal{S} = \mathcal{X}$, which selects

$$\boldsymbol{x}_n = \arg\max_{\boldsymbol{x}\in\mathcal{S}} \sum_{\boldsymbol{x}'\in\mathcal{A}} \mathrm{I}(f_{\boldsymbol{x}'}; y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1}) \tag{35}$$

and which simplifies in the GP setting to

$$\begin{aligned}
\boldsymbol{x}_n &= \arg\max_{\boldsymbol{x}\in\mathcal{S}} \frac{1}{2} \sum_{\boldsymbol{x}'\in\mathcal{A}} \log \left(\frac{\operatorname{Var}[f_{\boldsymbol{x}'} \mid \mathcal{D}_{n-1}]}{\operatorname{Var}[f_{\boldsymbol{x}'} \mid \mathcal{D}_{n-1}, y_{\boldsymbol{x}}]} \right) \\
&= \arg\min_{\boldsymbol{x}\in\mathcal{S}} \sum_{\boldsymbol{x}'\in\mathcal{A}} \log \operatorname{Var}[f_{\boldsymbol{x}'} \mid \mathcal{D}_{n-1}, y_{\boldsymbol{x}}] \\
&= \arg\min_{\boldsymbol{x}\in\mathcal{S}} \operatorname{tr} \log \operatorname{diag} \operatorname{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_{n-1}, y_{\boldsymbol{x}}].
\end{aligned} \tag{36}$$

Analogously to the derivation of Equation (33), this can also be expressed as

$$\boldsymbol{x}_n = \arg\min_{\boldsymbol{x}\in\mathcal{S}} |\operatorname{diag} \operatorname{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_{n-1}, y_{\boldsymbol{x}}]| \,. \tag{37}$$

Effectively, MM-ITL ignores the mutual interaction between points in $\mathcal{A}$. As can be seen from Equation (36) and as is also mentioned by MacKay (1992), MM-ITL is equivalent to VTL up to a different weighting of the points in $\mathcal{A}$: instead of minimizing the average posterior variance (as in VTL), MM-ITL minimizes the average posterior log-variance. Under the lens of principle (†), this can be seen as minimizing the average marginal entropy of predictions within the target space:

$$\boldsymbol{x}_n = \arg\min_{\boldsymbol{x}\in\mathcal{S}} \sum_{\boldsymbol{x}'\in\mathcal{A}} \mathrm{H}[f_{\boldsymbol{x}'} \mid \mathcal{D}_{n-1}, y_{\boldsymbol{x}}] \,.$$

We remark that MM-ITL coincides with EPIG (Smith et al., 2023, Appendix E.2).

Not a generalization of uncertainty sampling Unlike ITL, MM-ITL is not a generalization of uncertainty sampling. The reason is precisely that MM-ITL ignores the mutual interaction between points in $\mathcal{A}$. Consider the example where $\mathcal{X} = \mathcal{S} = \mathcal{A} = \{1, \ldots, 10\}$ where $\boldsymbol{f}_{1:9}$ are highly correlated while f_{10} is mostly independent of the other points. Visually, imagine a smooth function (i.e., under a Gaussian kernel) with points 1 through 9 close to each other and point 10 far away. Further, suppose that point 10 has a slightly larger marginal variance than the others. Then, MM-ITL would select one of the points $1:9$ since this leads to the largest reduction in the marginal (log-)variances (i.e., to a small posterior "uncertainty").[11] In contrast, ITL selects the point with the largest prior marginal variance (cf. Appendix C.1), point 10, since this leads to the largest reduction in entropy.[12]

Similarity to VTL Observe that the following decision rule is equivalent to VTL:

$$\boldsymbol{x}_n = \arg\max_{\boldsymbol{x}\in\mathcal{S}} \operatorname{tr} \operatorname{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_{n-1}] - \operatorname{tr} \operatorname{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_{n-1}, y_{\boldsymbol{x}}].$$

By Lemma C.38, for any $\boldsymbol{x} \in \mathcal{S}$, this objective value can be tightly lower- and upper-bounded (up to constant-factors) by

$$\begin{aligned}
&\sum_{\boldsymbol{x}'\in\mathcal{A}} \log\left(\frac{\operatorname{Var}[f_{\boldsymbol{x}'} \mid \mathcal{D}_{n-1}]}{\operatorname{Var}[f_{\boldsymbol{x}'} \mid \mathcal{D}_{n-1}, y_{\boldsymbol{x}}]}\right) \\
&= 2 \sum_{\boldsymbol{x}'\in\mathcal{A}} \mathrm{I}(f_{\boldsymbol{x}'}; y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1}) && \text{(see MM-ITL)} \\
&\overset{(i)}{=} -\sum_{\boldsymbol{x}'\in\mathcal{A}} \log\left(1 - \operatorname{Cor}[f_{\boldsymbol{x}'}, y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1}]^2\right) && (38)
\end{aligned}$$

where (i) is detailed in example 8.5.1 of Cover (1999). Thus, VTL and MM-ITL are closely related.

Experiments In our experiments with Gaussian processes from Figures 2 and 4, we observe that MM-ITL performs similarly to VTL.

Convergence of uncertainty We derive a convergence guarantee for MM-ITL which is analogous to the guarantees for ITL from Theorem C.12 and for VTL from Theorem C.13. We will assume for simplicity that Γ_n is monotonically decreasing in n (i.e., $\alpha_n \leq 1$).

Theorem D.1 (Convergence of uncertainty reduction of MM-ITL)**.** *Assume that Assumptions B.1 and B.2 are satisfied. Then for any $n \geq 1$, if $\Gamma_0 \geq \cdots \geq \Gamma_{n-1}$ and the sequence $\{\boldsymbol{x}_i\}_{i=1}^n$ is generated by* MM-ITL*, then*

$$\Gamma_{n-1} \leq \frac{1}{n} \sum_{\boldsymbol{x}'\in\mathcal{A}} \gamma_n(\{\boldsymbol{x}'\}; \mathcal{S}). \tag{39}$$

Proof. We have

$$\Gamma_{n-1} = \frac{1}{n} \sum_{i=0}^{n-1} \Gamma_{n-1}$$

[11] This is because the observation reduces uncertainty not just about the observed point itself.

[12] Because points $\boldsymbol{f}_{1:9}$ are highly correlated, $\mathrm{H}[\boldsymbol{f}_{1:9}]$ is already "small".

$$
\begin{aligned}
&\overset{(i)}{\leq} \frac{1}{n}\sum_{i=0}^{n-1}\Gamma_i \\
&= \frac{1}{n}\sum_{i=0}^{n-1}\max_{\boldsymbol{x}\in\mathcal{S}}\sum_{\boldsymbol{x}'\in\mathcal{A}}\mathrm{I}(f_{\boldsymbol{x}'};y_{\boldsymbol{x}}\mid\mathcal{D}_n) \\
&\overset{(ii)}{=} \frac{1}{n}\sum_{i=0}^{n-1}\sum_{\boldsymbol{x}'\in\mathcal{A}}\mathrm{I}(f_{\boldsymbol{x}'};y_{\boldsymbol{x}_{n+1}}\mid\mathcal{D}_n) \\
&\overset{(iii)}{=} \frac{1}{n}\sum_{\boldsymbol{x}'\in\mathcal{A}}\sum_{i=0}^{n-1}\mathrm{I}(f_{\boldsymbol{x}'};y_{\boldsymbol{x}_{n+1}}\mid\boldsymbol{y}_{\boldsymbol{x}_{1:n}}) \\
&\overset{(iv)}{=} \frac{1}{n}\sum_{\boldsymbol{x}'\in\mathcal{A}}\mathrm{I}\big(f_{\boldsymbol{x}'};\boldsymbol{y}_{\boldsymbol{x}_{1:n}}\big) \\
&\leq \frac{1}{n}\sum_{\boldsymbol{x}'\in\mathcal{A}}\max_{\substack{X\subseteq\mathcal{S}\\|X|=n}}\mathrm{I}(f_{\boldsymbol{x}'};\boldsymbol{y}_X) \\
&= \frac{1}{n}\sum_{\boldsymbol{x}'\in\mathcal{A}}\gamma_n(\{\boldsymbol{x}'\};\mathcal{S})
\end{aligned}
$$

where (i) follows by assumption; (ii) follows from the MM-ITL decision rule; (iii) uses that the posterior variance of Gaussians is independent of the realization and only depends on the *location* of observations; and (iv) uses the chain rule of mutual information. The remainder of the proof is analogous to the proof of Theorem C.12 (cf. Appendix C.4). □

Noting that

$$
\mathrm{I}(f_{\boldsymbol{x}'};y_{\boldsymbol{x}}\mid\mathcal{D}_{n-1}) \leq \sum_{\boldsymbol{x}'\in\mathcal{A}}\mathrm{I}(f_{\boldsymbol{x}'};y_{\boldsymbol{x}}\mid\mathcal{D}_{n-1})
$$

for any $n \geq 1$, $\boldsymbol{x} \in \mathcal{X}$, and $\boldsymbol{x}' \in \mathcal{A}$, Theorem 3.2 can be readily rederived for MM-ITL (cf. Lemmas C.14 and C.18). Hence, the posterior marginal variances of MM-ITL can be bounded uniformly in terms of Γ_n analogously to ITL.

D.4 Summary

We have seen that ITL, VTL, and MM-ITL can be seen as different interpretations of the same fundamental principle (†). If $|\mathcal{A}| = 1$ and correlations are non-negative, then all four decision rules are equivalent. ITL is the only decision rule that takes into account the mutual dependence between points in $\mathcal{A}$, and VTL and MM-ITL differ only in their weighting of the posterior marginal variances of points in $\mathcal{A}$.

E Stochastic Target Spaces

When the target space $\mathcal{A}$ is large, it may be computationally infeasible to compute the exact objective. A natural approach to address this issue is to approximate the target space by a smaller set of size K.

Target distributions Let $\mathcal{A} \subseteq \mathcal{X}$ be a (possibly continuous) target space, and let $\mathcal{P}_{\mathcal{A}}$ be a probability distribution supported on $\mathcal{A}$. In iteration n, a subset A_n of K points is sampled independently from $\mathcal{A}$ according to the distribution $\mathcal{P}_{\mathcal{A}}$ and the objective is computed on this subset. Formally, this amounts to a single-sample Monte Carlo approximation of

$$
\boldsymbol{x}_n \in \operatorname*{arg\,max}_{\boldsymbol{x}\in\mathcal{S}} \mathbb{E}_{A\overset{\text{iid}}{\sim}\mathcal{P}_{\mathcal{A}}}[\mathrm{I}(\boldsymbol{f}_A;y_{\boldsymbol{x}}\mid\mathcal{D}_{n-1})]. \tag{40}
$$

The convergence guarantees from Appendix C can be generalized to the setting of stochastic target spaces by estimating how often points "near" a specified prediction target $\boldsymbol{x} \in \mathcal{A}$ are sampled.

Definition E.1 (γ-ball at $\boldsymbol{x}$)**.** Given $\boldsymbol{x} \in \mathcal{A}$ and any $\gamma \geq 0$, we call the set

$$B_\gamma(\boldsymbol{x}) \stackrel{\text{def}}{=} \{\boldsymbol{x'} \in \mathcal{X} \mid \|\boldsymbol{x} - \boldsymbol{x'}\| \leq \gamma\}$$

the γ-ball at $\boldsymbol{x}$. Further, we call $\mathcal{P}_{\mathcal{A}}(B_\gamma(\boldsymbol{x}))$ the weight of that ball.

Proposition E.2. *Given any $n \geq 1, K \geq 1, \gamma > 0$, and $\boldsymbol{x} \in \mathcal{A}$, suppose that $B_\gamma(\boldsymbol{x})$ has weight $p > 0$. Assume that the* ITL *objective is L_I-Lipschitz continuous. Then, with probability at least $1 - \exp(-(1-p)n/(8K))$,*

$$\sigma_n^2(\boldsymbol{x}) \lesssim \eta_{\mathcal{S}}^2(\boldsymbol{x}) + CL_I\gamma\frac{\gamma_{k(n)}}{\sqrt{k(n)}}$$

where $k(n) \stackrel{\text{def}}{=} Kpn/2$.

Proof sketch. Let $Y_i \sim \mathrm{Binom}(K, p)$ denote the random variable counting the number of occurrences of a point from $B_\gamma(\boldsymbol{x})$ in A_i. Moreover, we write $X_i \stackrel{\text{def}}{=} \mathbb{1}\{B_\gamma(\boldsymbol{x}) \cap A_i \neq \emptyset\}$. Note that

$$\nu \stackrel{\text{def}}{=} \mathbb{E}X_i = \mathbb{P}(B_\gamma(\boldsymbol{x}) \cap A_i \neq \emptyset) = 1 - \mathbb{P}(Y_i = 0) = 1 - (1-p)^K \approx Kp$$

where the approximation stems from a first-order truncation of the Bernoulli series. Let $X \stackrel{\text{def}}{=} \sum_{i=1}^n X_i$ with $\mathbb{E}X = n\nu \approx Kpn$.

Using the assumed Lipschitz-continuity of the objective, we know that $\mathrm{I}(\boldsymbol{f}_{A'}; y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1}) \leq L_I\gamma\mathrm{I}(\boldsymbol{f}_A; y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1})$ where $A' \stackrel{\text{def}}{=} (A \setminus \{\boldsymbol{x}_\gamma\}) \cup \{\boldsymbol{x}\}$ and $\boldsymbol{x}_\gamma$ is the point from the γ-ball at $\boldsymbol{x}$. The bound then follows analogously to Theorem 3.2.

Finally, by Chernoff's bound, at least $Kpn/2$ iterations contain a point from $B_\gamma(\boldsymbol{x})$ with probability at least $1 - \exp(-Kpn/8)$. □

This strategy can also be used to generalize the VTL and MM-ITL objectives to stochastic target spaces.

F Closed-form Decision Rules

Below, we list the closed-form expressions for the ITL and VTL objectives. In the following, k_n denotes the kernel conditional on $\mathcal{D}_n$.

ITL

$$\begin{aligned}\mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1}) &= \frac{1}{2}\log\left(\frac{\mathrm{Var}[y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1}]}{\mathrm{Var}[y_{\boldsymbol{x}} \mid \boldsymbol{f}_{\mathcal{A}}, \mathcal{D}_{n-1}]}\right) \\ &= \frac{1}{2}\log\left(\frac{k_{n-1}(\boldsymbol{x},\boldsymbol{x}) + \rho^2}{\hat{k}_{n-1}(\boldsymbol{x},\boldsymbol{x}) + \rho^2}\right)\end{aligned} \tag{41}$$

where $\hat{k}_n(\boldsymbol{x},\boldsymbol{x}) = k_n(\boldsymbol{x},\boldsymbol{x}) - \boldsymbol{k}_n(\boldsymbol{x},\mathcal{A})\boldsymbol{K}_n(\mathcal{A},\mathcal{A})^{-1}\boldsymbol{k}_n(\mathcal{A},\boldsymbol{x})$.

VTL

$$\mathrm{tr}\ \mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_{n-1}, y_{\boldsymbol{x}}] = \sum_{\boldsymbol{x'} \in \mathcal{A}}\left(k_{n-1}(\boldsymbol{x'},\boldsymbol{x'}) - \frac{k_{n-1}(\boldsymbol{x},\boldsymbol{x'})^2}{k_{n-1}(\boldsymbol{x},\boldsymbol{x}) + \rho^2}\right).$$

G Computational Complexity

Evaluating the acquisition function of ITL in round n requires computing for each $\boldsymbol{x} \in \mathcal{S}$,

$$\begin{aligned}&\mathrm{I}(\boldsymbol{f}_{\mathcal{A}}; y_{\boldsymbol{x}} \mid \mathcal{D}_n) \\ &= \frac{1}{2}\log\left(\frac{|\mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_n]|}{|\mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid y_{\boldsymbol{x}}, \mathcal{D}_n]|}\right) && \text{(forward)} \\ &= \frac{1}{2}\log\left(\frac{\mathrm{Var}[y_{\boldsymbol{x}} \mid \mathcal{D}_n]}{\mathrm{Var}[y_{\boldsymbol{x}} \mid \boldsymbol{f}_{\mathcal{A}}, \mathcal{D}_n]}\right) && \text{(backward).}\end{aligned}$$

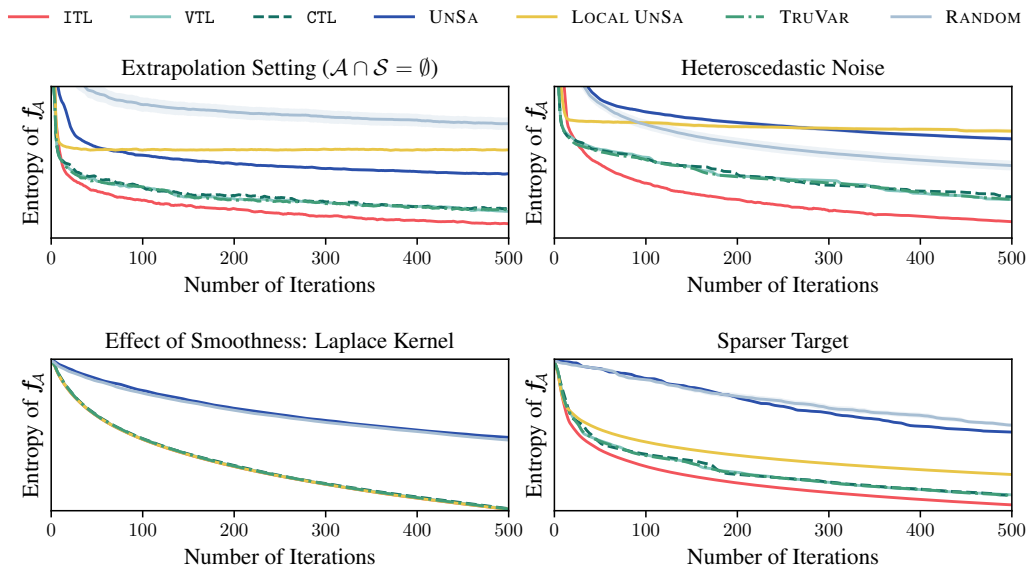

Figure 4: Additional GP experiments

Let $|\mathcal{S}| = m$ and $|\mathcal{A}| = k$. Then, the forward method has complexity $O\big(m \cdot k^3\big)$. For the backward method, observe that the variances are scalar and the covariance matrix $\mathrm{Var}[\boldsymbol{f}_{\mathcal{A}} \mid \mathcal{D}_n]$ only has to be inverted once for all points $\boldsymbol{x}$. Thus, the backward method has complexity $O\big(k^3 + m\big)$.

When the size m of $\mathcal{S}$ is relatively small (and hence, all points in $\mathcal{S}$ can be considered during each iteration of the algorithm), GP inference corresponds simply to computing conditional distributions of a multivariate Gaussian. The performance can therefore be improved by keeping track of the full posterior distribution over $\boldsymbol{f}_{\mathcal{S}}$ of size $O\big(m^2\big)$ and conditioning on the latest observation during each iteration of the algorithm. In this case, after each observation the posterior can be updated at a cost of $O\big(m^2\big)$ which does not grow with the time n, unlike classical GP inference.

Overall, when m is small, the computational complexity of ITL is $O\big(k^3 + m^2\big)$. When m is large (or possibly infinite) and a subset of $\tilde{m}$ points is considered in a given iteration, the computational complexity of ITL is $O\big(k^3 + \tilde{m} \cdot n^3\big)$, neglecting the complexity of selecting the $\tilde{m}$ candidate points. In the latter case, the computational cost of ITL is dominated by the cost of GP inference.

Khanna et al. (2017) discuss distributed and stochastic approximations of greedy algorithms to (weakly) submodular problems that are also applicable to ITL.

H Additional GP Experiments & Details

We use homoscedastic Gaussian noise with standard deviation $\rho = 0.1$ and a discretization of $\mathcal{X} = [-3, 3]^2$ of size 2 500. Uncertainty bands correspond to one standard error over 10 random seeds.

Additional experiments Figure 4 includes the following additional experiments:

1. *Extrapolation Setting ($\mathcal{A} \cap \mathcal{S} = \emptyset$)*: Right experiment from Figure 2 (A) under the Gaussian kernel. ITL has a similar advantage as in the setting shown in Figure 2 (B).
2. *Heteroscedastic Noise*: Left experiment from Figure 2 (A) under the Gaussian kernel with heteroscedastic Gaussian noise

$$\rho(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \boldsymbol{x} \in [-\frac{1}{2}, \frac{1}{2}]^2 \\ 0.1 & \text{otherwise} \end{cases}.$$

   If observation noise is heteroscedastic, in considering *posterior* rather than *prior* uncertainty, ITL avoids points with high aleatoric uncertainty, which accelerates learning.
3. *Effect of Smoothness*: Top experiment from Figure 2 (B) under the Laplace kernel. All algorithms except for US and RANDOM perform equally well. This validates our claims

from Section 3.3: in the extreme non-smooth case of a Laplace kernel and $\mathcal{A} \subseteq \mathcal{S}$, points outside $\mathcal{A}$ do not provide any additional information, and ITL and "local" UNSA coincide.

4. *Sparser Target*: Top experiment from Figure 2 (B) under the Gaussian kernel, but with domain extended to $\mathcal{X} = [-10, 10]^2$.

Hyperparameters of TRUVAR As suggested by Bogunovic et al. (2016), we use $\tilde{\eta}^2_{(1)} = 1$, $r = 0.1$, and $\delta = 0$ (even though the theory only holds for $\delta > 0$). The TRUVAR baseline only applies when $\mathcal{A} \subseteq \mathcal{S}$ (cf. Section 5).

Smoothing to reduce numerical noise Applied running average with window 5 to entropy curves of Figures 2 and 4 to smoothen out numerical noise.

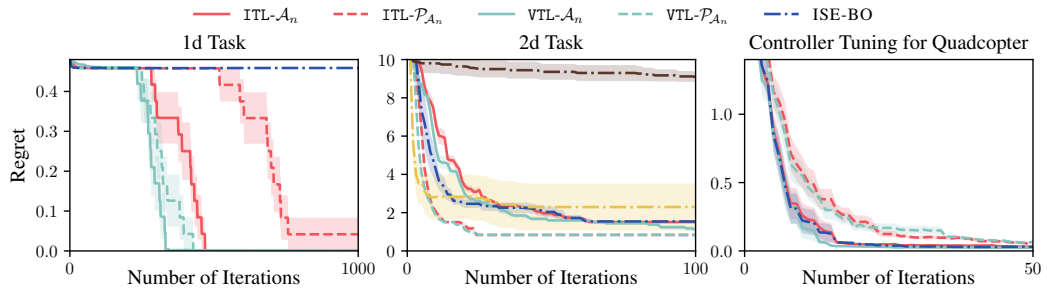

Figure 5: We perform the tasks of Figure 3 using Thompson sampling to evaluate the stochastic target space $\mathcal{P}_{\mathcal{A}_n}$. We additionally compare to GoOSE (cf. Appendix I.2.3) and ISE-BO (cf. Appendix I.2.4).

I Additional Safe BO Experiments & Details

In Appendix I.1, we discuss the use of stochastic target spaces in the safe BO setting. We provide a comprehensive overview of prior works in Appendix I.2 and an additional experiment highlighting that ITL, unlike SafeOpt, is able to "jump past local barriers" in Appendix I.3. In Appendix I.4, we provide details on the experiments from Figure 3.

I.1 A More Exploitative Stochastic Target Space

Alternatively to the target space $\mathcal{A}_n$ which comprises all potentially optimal points, we evaluate the stochastic target space

$$\mathcal{P}_{\mathcal{A}_n}(\cdot) = \mathbb{P}(\operatorname*{arg\,max}_{\boldsymbol{x}\in\mathcal{X}:g(\boldsymbol{x})\geq 0} f(\boldsymbol{x}) = \cdot \mid \mathcal{D}_n) \tag{42}$$

which effectively weights points in $\mathcal{A}_n$ according to how likely they are to be the safe optimum, and is therefore more exploitative than the uniformly-weighted target space discussed so far. Samples from $\mathcal{P}_{\mathcal{A}_n}$ can be obtained efficiently via Thompson sampling (Thompson, 1933; Russo et al., 2018). Observe that $\mathcal{P}_{\mathcal{A}_n}$ is supported precisely on the set of potential maximizers $\mathcal{A}_n$. We provide a formal analysis of stochastic target spaces in Appendix E. Whether transductive active learning with $\mathcal{A}_n$ or $\mathcal{P}_{\mathcal{A}_n}$ performs better is task-dependent, as we will see in the following.

Note that performing ITL with this target space is analogous to output-space entropy search (Wang & Jegelka, 2017). Samples from $\mathcal{P}_{\mathcal{A}_n}$ can be obtained via Thompson sampling (Thompson, 1933; Russo et al., 2018). That is, in iteration $n+1$, we sample $K \in \mathbb{N}$ independent functions $f^{(j)} \sim f \mid \mathcal{D}_n$ from the posterior distribution and select K points $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(K)}$ which are a safe maximum of $f^{(1)}, \dots, f^{(K)}$, respectively.

Experiments In Figure 5, we contrast the performance of ITL with $\mathcal{P}_{\mathcal{A}_n}$ to the performance of ITL with the exact target space $\mathcal{A}_n$. We observe that their relative performance is instance dependent: in tasks that require more difficult expansion, ITL with $\mathcal{A}_n$ converges faster, whereas in simpler tasks (such as the 2d experiment), ITL with $\mathcal{P}_{\mathcal{A}_n}$ converges faster. We compare against the GoOSE algorithm (Turchetta et al., 2019) which is a heuristic extension of SafeOpt that explores more greedily in directions of (assumed) high reward (cf. Appendix I.2.3). GoOSE suffers from the same limitations as SafeOpt, which were highlighted in Section 4, and additionally is limited by its heuristic approach to expansion which fails in the 1d task and safe controller tuning task. Analogously to our experiments with SafeOpt, we also compare against Oracle GoOSE which has oracle knowledge of the true Lipschitz constants.

The different behaviors of ITL with $\mathcal{A}_n$ and $\mathcal{P}_{\mathcal{A}_n}$, respectively, as well as SafeOpt and GoOSE are illustrated in Figure 6. We observe that ITL with $\mathcal{A}_n$ and SafeOpt expand the safe set more "uniformly" since the set of potential maximizers encircles the true safe set.[13] Intuitively, this is because the set of potential maximizers *conservatively* captures migh points might be safe and

[13]This is because typically, there will always remain points in $\widehat{\mathcal{S}}_n \setminus \mathcal{S}_n$ of which the safety cannot be fully determined, and since, they cannot be observed, it can also not be ruled out that they have high objective value.

optimal. In contrast, ITL with $\mathcal{P}_{\mathcal{A}_n}$ and GoOSE focus exploration and expansion in those regions where the objective is likely to be high.

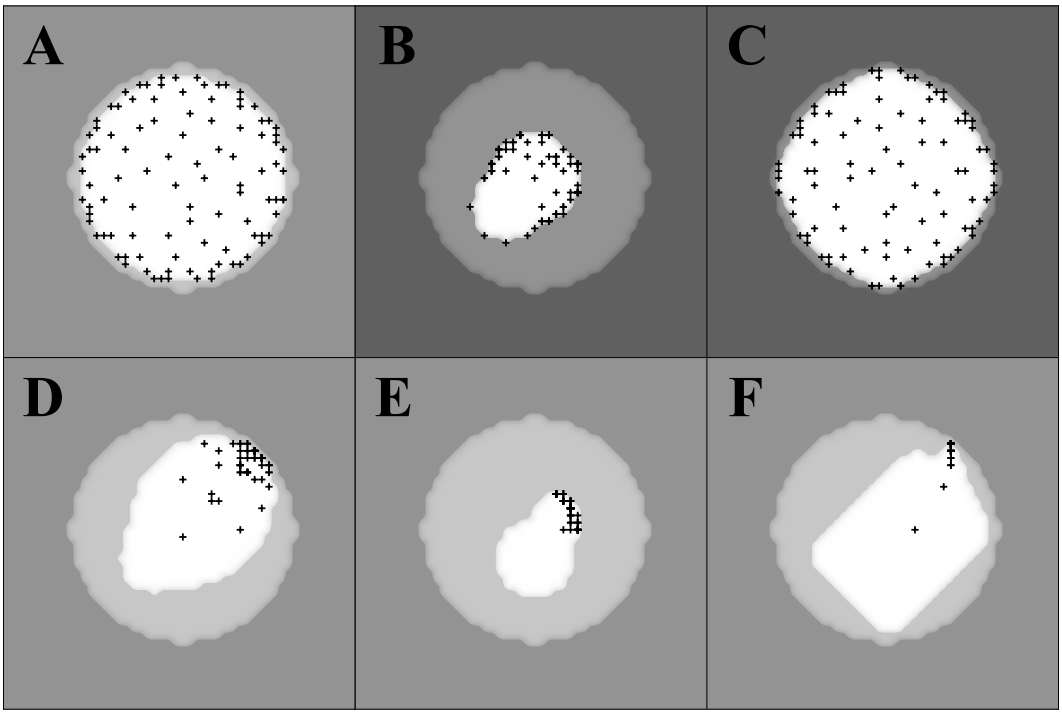

Figure 6: The first 100 samples of **(A)** ITL with $\mathcal{A}_n$, **(B)** SafeOpt, **(C)** Oracle SafeOpt, **(D)** ITL with $\mathcal{P}_{\mathcal{A}_n}$, **(E)** GoOSE, **(F)** Oracle GoOSE. The white region denotes the pessimistic safe set $\mathcal{S}_{100}$, the light gray region denotes the true safe set $\mathcal{S}^\star$ (i.e., the "island"), and the darker gray regions denotes unsafe points (i.e., the "ocean").

I.2 Detailed Comparison with Prior Works

The most widely used method for Safe BO is SafeOpt (Sui et al., 2015; Berkenkamp et al., 2021) which keeps track of separate candidate sets for expansion and exploration and uses UnSa to pick one of the candidates in each round. Treating expansion and exploration separately, sampling is directed towards expansion in *all* directions — even those that are known to be suboptimal. The safe set is expanded based on a Lipschitz constant of $g^\star$, which is assumed to be known. In most real-world settings, this constant is unknown and has to be estimated using the GP. This estimate is generally conservative and results in suboptimal performance. To this end, Berkenkamp et al. (2016) proposed Heuristic SafeOpt which relies solely on the confidence intervals of g to expand the safe set, but lacks convergence guarantees. More recently, Bottero et al. (2022) proposed ISE which queries parameters from $\mathcal{S}_n$ that yield the most "information" about the safety of another parameter in $\mathcal{X}$. Hence, ISE focuses solely on the expansion of the safe set $\mathcal{S}_n$ and does not take into account the objective f. In practice, this can lead to significantly worse performance on the simplest of problems (cf. Figure 3). In contrast, ITL balances expansion of and exploration within the safe set. Furthermore, ISE does not have known convergence guarantees of the kind of Theorem 4.1. In parallel independent work, Bottero et al. (2024) proposed a combination of ISE and max-value entropy search (Wang & Jegelka, 2017) for which they derive a similar guarantee to Theorem 4.1.[14] Similar to SafeOpt, their method aims to expand the safe set in all directions including those that are known to be suboptimal. In contrast, ITL directs expansion only towards potentially optimal regions.

In the 1d task and quadcopter experiment (cf. Figure 3), we observe that SafeOpt and even Oracle SafeOpt converge significantly slower than ITL to the safe optima. We believe this is due to their conservative Lipschitz-continuity/global smoothness-based expansion, as opposed to ITL's expansion,

[14]We provide an empirical evaluation in Appendix I.2.4.

which adapts to the local smoothness of the constraints. HEURISTIC SAFEOPT, which does not rely on the Lipschitz constant for expansion, does not efficiently expand the safe set due to its heuristic that only considers single-step expansion. This is especially the case for the 1d task. Furthermore, in the 2d task, we notice the suboptimality of ISE since it does not take into account the objective, and purely aims to expand the safe set. ITL, on the other hand, balances expansion and exploration.

I.2.1 SAFEOPT

SAFEOPT (Sui et al., 2015; Berkenkamp et al., 2021) is a well-known algorithm for Safe BO.

Lipschitz-based expansion SAFEOPT expands the set of known-to-be safe points by assuming knowledge of an upper bound L_i to the Lipschitz constant of the unknown constraints $g_i^\star$.[15] In each iteration, the (pessimistic) safe set $\mathcal{S}_n$ is updated to include all points which can be reached safely (with respect to the Lipschitz continuity) from a known-to-be-safe point $\boldsymbol{x} \in \mathcal{S}_n$. Formally,

$$\mathcal{S}_n^{\textsc{SafeOpt}} \stackrel{\text{def}}{=} \bigcup_{\boldsymbol{x} \in \mathcal{S}_{n-1}^{\textsc{SafeOpt}}} \{\boldsymbol{x}' \in \mathcal{X} \mid l_{n,i}(\boldsymbol{x}) - L_i \|\boldsymbol{x} - \boldsymbol{x}'\|_2 \geq 0 \text{ for all } i \in \mathcal{I}_s\}. \tag{43}$$

The expansion of the safe set is illustrated in Figure 7.

We remark two main limitations of this approach. First, the Lipschitz constant is an additional safety critical hyperparameter of the algorithm, which is typically not known. The RKHS assumption (cf. Assumption C.25) induces an assumption on the Lipschitz continuity, however, the worst-case a-priori Lipschitz constant is typically very large, and prohibitive for expansion. Second, the Lipschitz constant is global property of the unknown function, meaning that it does not adapt to the local smoothness. For example, a constraint may be "flat" in one direction (permitting straightforward expansion) and "steep" in another direction (requiring slow expansion). Furthermore, the Lipschitz constant is constant over time, whereas ITL is able to adapt to the local smoothness and reduce the (induced) Lipschitz constant over time.

Undirected expansion SAFEOPT addresses the trade-off between expansion and exploration by focusing learning on two different sets. First, the set of *maximizers*

$$\mathcal{M}_n^{\textsc{SafeOpt}} \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \mathcal{S}_n^{\textsc{SafeOpt}} \mid u_{n,f}(\boldsymbol{x}) \geq \max_{\boldsymbol{x}' \in \mathcal{S}_n^{\textsc{SafeOpt}}} l_{n,f}(\boldsymbol{x})\}$$

which contains all *known-to-be-safe* points which are potentially optimal. Note that if $\mathcal{S}_n^{\textsc{SafeOpt}} = \mathcal{S}_n$ then $\mathcal{M}_n^{\textsc{SafeOpt}} \subseteq \mathcal{A}_n$ since $\mathcal{A}_n$ contains points which are potentially optimal and potentially safe *but possibly unsafe*.

To facilitate expansion, for each point $\boldsymbol{x} \in \mathcal{S}_n$, the algorithm considers a set of *expanding points*

$$\mathcal{F}_n^{\textsc{SafeOpt}}(\boldsymbol{x}) \stackrel{\text{def}}{=} \{\boldsymbol{x}' \in \mathcal{X} \setminus \mathcal{S}_n^{\textsc{SafeOpt}} \mid u_{n,i}(\boldsymbol{x}) - L_i \|\boldsymbol{x} - \boldsymbol{x}'\|_2 \geq 0 \text{ for all } i \in \mathcal{I}_s\}$$

A point is expanding if it is unsafe initially and can be (optimistically) deduced as safe by observing $\boldsymbol{x}$. The set of *expanders* corresponds to all known-to-be-safe points which optimistically lead to expansion of the safe set:

$$\mathcal{G}_n^{\textsc{SafeOpt}} \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \mathcal{S}_n \mid |\mathcal{F}_n(\boldsymbol{x})| > 0\}.$$

That is, an expander is a safe point $\boldsymbol{x}$ which is "close" to at least one expanding point $\boldsymbol{x}'$. Observe that here, we start with a safe $\boldsymbol{x}$ and then find a close and potentially safe $\boldsymbol{x}'$ using the Lipschitz-property of the constraint function. Thus, the set of expanding points is inherently limited by the assumed Lipschitzness (cf. Figure 7), and generally a subset of the potential expanders $\mathcal{E}_n$ (cf. Equation (26)):

Lemma I.1. *For any $n \geq 0$, if $\mathcal{S}_n^{\textsc{SafeOpt}} = \mathcal{S}_n$ then*

$$\bigcup_{\boldsymbol{x} \in \mathcal{S}_n} \mathcal{F}_n^{\textsc{SafeOpt}}(\boldsymbol{x}) \subseteq \mathcal{E}_n.$$

[15] Recall that due to the assumption that $\|g_i^\star\|_k < \infty$, $g_i^\star$ is indeed Lipschitz continuous.

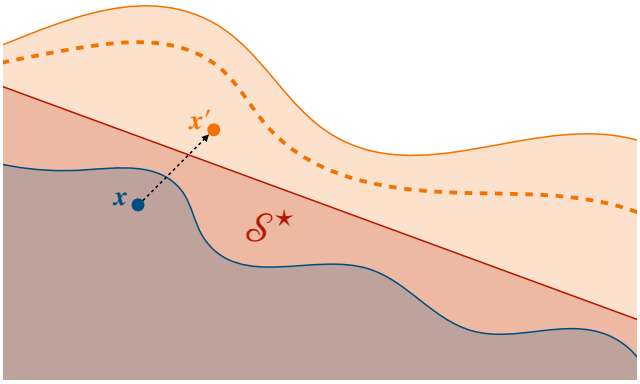

Figure 7: Illustration of the expansion of the safe set à la SAFEOPT. Here, the blue region denotes the pessimistic safe set $\mathcal{S}$, the red region denotes the true safe set $\mathcal{S}^\star$, and the orange region denotes the optimistic safe set $\widehat{\mathcal{S}}$. Whereas ITL learns about the point $\boldsymbol{x'}$ *directly*, SAFEOPT expands the safe set using the reduction of uncertainty at $\boldsymbol{x}$, and then extrapolating using the Lipschitz constant (cf. Equation (43)). The dashed orange line denotes the expanding points of SAFEOPT which under-approximate the optimistic safe set of ITL (cf. Lemma I.1). Thus, ITL may even learn about points in $\widehat{\mathcal{S}}$ which are "out of reach" for SAFEOPT.

Proof. Without loss of generality, we consider the case where $\mathcal{I}_s = \{i\}$. We have

$$\mathcal{E}_n = \widehat{\mathcal{S}}_n \setminus \mathcal{S}_n = \{\boldsymbol{x} \in \mathcal{X} \setminus \mathcal{S}_n \mid u_{n,i}(\boldsymbol{x}) \geq 0\}.$$

The result follows directly by observing that $L_i \|\boldsymbol{x} - \boldsymbol{x'}\|_2 \geq 0$. □

SAFEOPT then selects $\boldsymbol{x}_{n+1}$ according to uncertainty sampling *within* the maximizers and expanders: $\mathcal{M}_n^{\text{SafeOpt}} \cup \mathcal{G}_n^{\text{SafeOpt}}$. We remark that due to the separate handling of expansion and exploration, SAFEOPT expands the safe set in *all* directions — even those that are known to be suboptimal. In contrast, ITL only expands the safe set in directions that are potentially optimal by balancing expansion and exploration through the single set of potential maximizers $\mathcal{A}_n$.

Based on uncertainty sampling As mentioned in the previous paragraph, SAFEOPT selects as next point the maximizer/expander with the largest prior uncertainty.[16] In contrast, ITL selects the point within $\mathcal{S}_n$ which minimizes the posterior uncertainty within $\mathcal{A}_n$. Note that the two approaches are not identical as typically $\mathcal{M}_n^{\text{SafeOpt}} \cup \mathcal{G}_n^{\text{SafeOpt}} \subset \mathcal{S}_n^{\text{SafeOpt}}$ and $\mathcal{A}_n \not\supseteq \mathcal{S}_n$.

We show empirically in Section 3.3 that depending on the kernel choice (i.e., the smoothness assumptions), uncertainty sampling within a given target space neglects higher-order information that can be attained by sampling outside the set. This can be seen even more clearly when considering linear functions, in which case points outside the maximizers and expanders can be equally informative as points inside.

Finally, note that the set of expanders is constructed "greedily", i.e., only considering *single-step* expansion. This is necessitated as the inference of safety is based on single reference points. Instead, ITL directly quantifies the information gained towards the points of interest without considering intermediate reference points.

Requires homoscedastic noise SAFEOPT imposes a homoscedasticity assumption on the noise which is an artifact of the analysis of uncertainty sampling. It is well known that in the presence of heteroscedastic noise, one has to distinguish epistemic and aleatoric uncertainty. Uncertainty sampling fails because it may continuously sample a high variance point where the variance is dominated by aleatoric uncertainty, potentially missing out on reducing epistemic uncertainty at points with small aleatoric uncertainty. In contrast, maximizing mutual information naturally takes

[16] The use of uncertainty sampling for safe sequential decision-making goes back to Schreiter et al. (2015) and Sui et al. (2015).

into account the two sources of uncertainty, preferring those points where epistemic uncertainty is large and aleatoric uncertainty is small (cf. Appendix C.1).

Suboptimal reachable safe set Sui et al. (2015) and Berkenkamp et al. (2021) show that SAFEOPT converges to the optimum within the closure $\bar{\mathcal{R}}_\epsilon^{\text{SafeOpt}}(\mathcal{S}_0)$ of

$$\mathcal{R}_\epsilon^{\text{SafeOpt}}(\mathcal{S}) \stackrel{\text{def}}{=} \mathcal{S} \cup \{\boldsymbol{x} \in \mathcal{X} \mid \exists \boldsymbol{x'} \in \mathcal{S} \text{ such that } f_i^\star(\boldsymbol{x'}) - (L_i \|\boldsymbol{x} - \boldsymbol{x'}\|_2 + \epsilon) \geq 0 \text{ for all } i \in \mathcal{I}_s\}.$$

Note that analogously to the expansion of the safe set, the "expansion" of the reachable safe set is based on "inferring safety" through a reference point in $\mathcal{S}$ and using Lipschitz continuity. This is opposed to the reachable safe set of ITL (cf. Definition C.29).

We remark that under the additional assumption that a Lipschitz constant is known, ITL can easily be extended to expand its safe set based on the kernel *and* the Lipschitz constant, resulting in a strictly larger reachable safe set than SAFEOPT. We leave the concrete formalization of this extension to future work. Moreover, we do not evaluate this extension in our experiments, as we observe that even without the additional assumption of a Lipschitz constant, ITL outperforms SAFEOPT in practice.

I.2.2 HEURISTIC SAFEOPT

Berkenkamp et al. (2016) also implement a heuristic variant of SAFEOPT which does not assume a known Lipschitz constant. This heuristic variant uses the same (pessimistic) safe sets $\mathcal{S}_n$ as ITL. The set of maximizers is identical to SAFEOPT. As expanders, the heuristic variant considers all safe points $\boldsymbol{x} \in \mathcal{S}_n$ that if $\boldsymbol{x}$ were to be observed next with value $\boldsymbol{u}_n(\boldsymbol{x})$ lead to $|\mathcal{S}_{n+1}| > |\mathcal{S}_n|$. We refer to this set as $\mathcal{G}_n^{\text{H-SafeOpt}}$. The next point is then selected by uncertainty sampling within $\mathcal{M}_n^{\text{SafeOpt}} \cup \mathcal{G}_n^{\text{H-SafeOpt}}$.

The heuristic variant shares some properties with SAFEOPT, such that it is based on uncertainty sampling, not adapting to heteroscedastic noise, and separate notions of maximizers and expanders (leading to an "undirected" expansion of the safe set). Note that there are no known convergence guarantees for heuristic SAFEOPT. Importantly, note that similar to SAFEOPT the set of expanders is constructed "greedily", and in particular, does only take into account *single-step* expansion. In contrast, an objective such as ITL which quantifies the "information gained towards expansion" also actively seeks out *multi-step* expansion.

I.2.3 GOOSE

To address the "undirected" expansion of SAFEOPT discussed in the previous section, Turchetta et al. (2019) proposed *goal-oriented safe exploration* (GOOSE). GOOSE extends any unsafe BO algorithm (which we subsequently call an oracle) to the safe setting. In our experiments, we evaluate GOOSE-UCB which uses UCB as oracle and which is also the variant studied by Turchetta et al. (2019). In the following, we assume for ease of notation that $\mathcal{I}_s = \{c\}$.

Given the oracle proposal $\boldsymbol{x}^\star$, GOOSE first determines whether $\boldsymbol{x}^\star$ is safe. If $\boldsymbol{x}^\star$ is safe, $\boldsymbol{x}^\star$ is queried next. Otherwise, GOOSE first learns about the safety of $\boldsymbol{x}^\star$ by querying "expansionist" points until the oracle's proposal is determined to be either safe or unsafe.

GOOSE expands the safe set identically to SAFEOPT according to Equation (43). In the context of GOOSE, $\mathcal{S}_n^{\text{SafeOpt}}$ is called the *pessimistic safe set*. To determine that a point cannot be deduced as safe, GOOSE also keeps track of a Lipschitz-based *optimistic safe set*:

$$\widehat{\mathcal{S}}_{n,\epsilon}^{\text{GoOSE}} \stackrel{\text{def}}{=} \bigcup_{\boldsymbol{x} \in \mathcal{S}_{n-1}^{\text{SafeOpt}}} \{\boldsymbol{x'} \in \mathcal{X} \mid u_{n,c}(\boldsymbol{x}) - L_c \|\boldsymbol{x} - \boldsymbol{x'}\|_2 - \epsilon \geq 0\}.$$

We summarize the algorithm in Algorithm 1 where we denote by $\mathcal{O}(\mathcal{X})$ the oracle proposal over the domain $\mathcal{X}$.

It remains to discuss the heuristic used to select the "expansionist" points. GOOSE considers all points $\boldsymbol{x} \in \mathcal{S}_n^{\text{SafeOpt}}$ with confidence bands of size larger than the accuracy ϵ, i.e.,

$$\mathcal{W}_{n,\epsilon}^{\text{GoOSE}} \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \mathcal{S}_n^{\text{SafeOpt}} \mid u_{n,c}(\boldsymbol{x}) - l_{n,c}(\boldsymbol{x}) > \epsilon\}.$$

Algorithm 1 GoOSE

Given: Lipschitz constant L_c, prior model $\{f, g_c\}$, oracle $\mathcal{O}$, and precision ϵ
Set initial safe set $\mathcal{S}_0^{\text{SafeOpt}}$ based on prior
$\widehat{\mathcal{S}}_{n,\epsilon}^{\text{GoOSE}} \leftarrow \mathcal{X}$
$n \leftarrow 0$
for k from 1 to ∞ **do**
  $\boldsymbol{x}_k^\star \leftarrow \mathcal{O}(\widehat{\mathcal{S}}_{n,\epsilon}^{\text{GoOSE}})$
  while $\boldsymbol{x}_k^\star \notin \mathcal{S}_n^{\text{SafeOpt}}$ **do**
    Observe "expansionist" point $\boldsymbol{x}_{n+1}$, set $n \leftarrow n+1$, and update model and safe sets
  end while
  Observe $\boldsymbol{x}_k^\star$, set $n \leftarrow n+1$, and update model and safe sets
end for

Which of the points in this set is evaluated depends on a set of learning targets $\mathcal{A}_{n,\epsilon}^{\text{GoOSE}} \stackrel{\text{def}}{=} \widehat{\mathcal{S}}_{n,\epsilon}^{\text{GoOSE}} \setminus \mathcal{S}_n^{\text{SafeOpt}}$ akin to the "potential expanders" $\mathcal{E}_n$ (cf. Equation (26)), to each of which we assign a priority $h(\boldsymbol{x})$. When $h(\boldsymbol{x})$ is large, this indicates that the algorithm is prioritizing to determine whether $\boldsymbol{x}$ is safe. We use as heuristic the negative ℓ_1-distance between $\boldsymbol{x}$ and $\boldsymbol{x}^\star$. GoOSE then considers the set of *potential immediate expanders*

$$\mathcal{G}_{n,\epsilon}^{\text{GoOSE}}(\alpha) \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \mathcal{W}_{n,\epsilon}^{\text{GoOSE}} \mid \exists \boldsymbol{x'} \in \mathcal{A}_{n,\epsilon}^{\text{GoOSE}} \text{ with priority } \alpha \text{ such that } u_{n,c}(\boldsymbol{x}) - L_c\|\boldsymbol{x} - \boldsymbol{x'}\|_2 \geq 0\}.$$

The "expansionist" point selected by GoOSE is then any point in $\mathcal{G}_{n,\epsilon}^{\text{GoOSE}}(\alpha^\star)$ where $\alpha^\star$ denotes the largest priority such that $|\mathcal{G}_{n,\epsilon}^{\text{GoOSE}}(\alpha^\star)| > 0$.

We observe empirically that the sample complexity of GoOSE is not always better than that of SafeOpt. Notably, the expansion of the safe set is based on a "greedy" heuristic. Moreover, determining whether a single oracle proposal $\boldsymbol{x}^\star$ is safe may take significant time. Consider the (realistic) example where the prior is uniform, and UCB proposes a point which is far away from the safe set and suboptimal. GoOSE will typically attempt to derive the safety of the proposed point until the uncertainty at *all* points within $\mathcal{S}_0^{\text{SafeOpt}}$ is reduced to ϵ.[17] Thus, GoOSE can "waste" a significant number of samples, aiming to expand the safe set towards a known-to-be suboptimal point. In larger state spaces, due to the greedy nature of the expansion strategy, this can lead to GoOSE being effectively stuck at a suboptimal point for a significant number of rounds.

I.2.4 ISE and ISE-BO

Recently, Bottero et al. (2022) proposed an information-theoretic approach to efficiently expand the safe set which they call *information-theoretic safe exploration* (ISE). Specifically, they choose the next action $\boldsymbol{x}_n$ by approximating

$$\arg\max_{\boldsymbol{x} \in \mathcal{S}_{n-1}} \underbrace{\max_{\boldsymbol{x'} \in \mathcal{X}} \mathrm{I}(\mathbb{1}\{g_{\boldsymbol{x'}} \geq 0\}; y_{\boldsymbol{x}} \mid \mathcal{D}_{n-1})}_{\alpha^{\text{ISE}}(\boldsymbol{x})}. \tag{ISE}$$

In a parallel independent work, Bottero et al. (2024) extended ISE to the Safe BO problem where they propose to choose $\boldsymbol{x}_n$ according to

$$\arg\max_{\boldsymbol{x} \in \mathcal{S}_{n-1}} \max\{\alpha^{\text{ISE}}(\boldsymbol{x}), \alpha^{\text{MES}}(\boldsymbol{x})\} \tag{ISE-BO}$$

where α^{MES} denotes the acquisition function of max-value entropy search (Wang & Jegelka, 2017). Similarly to SafeOpt, ISE-BO treats expansion and exploration separately, which leads to "undirected" expansion of the safe set. That is, the safe set is expanded in all directions, even those that are known to be suboptimal. In contrast, ITL balances expansion and exploration through the single set of potential maximizers $\mathcal{A}_n$. With a stochastic target space, ITL generalizes max-value entropy search (cf. Appendix I.1).

We evaluate ISE-BO in Figure 5 and observe that it does not outperform ITL and VTL in any of the tasks, while performing poorly in the 1d task and suboptimally in the 2d task.

[17] This is because the proposed point typically remains in the optimistic safe set when it is sufficiently far away from the pessimistic safe set.

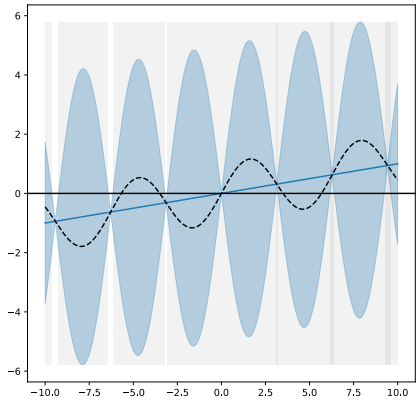

Figure 8: The ground truth $f^\star$ is shown as the dashed black line. The solid black line denotes the constraint boundary. The GP prior is given by a linear kernel with $\sin$-transform and mean $0.1x$. The light gray region denotes the initial optimistic safe set $\widehat{\mathcal{S}}_0$ and the dark gray region denotes the initial pessimistic safe set $\mathcal{S}_0$.

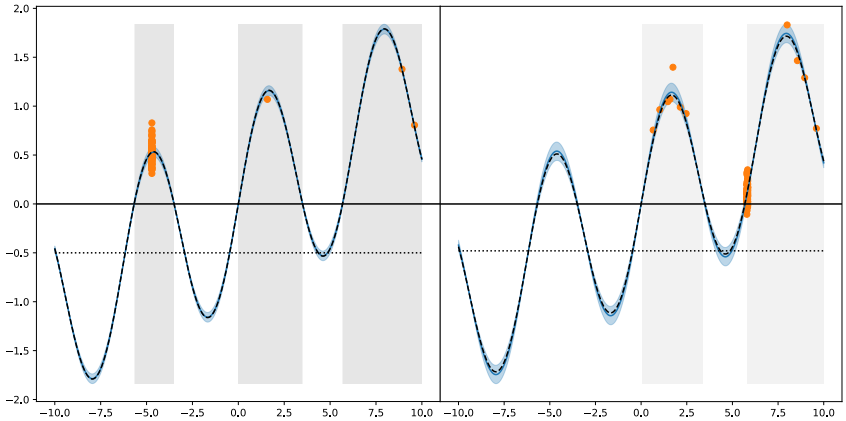

Figure 9: First 100 samples of ITL using the potential expanders $\mathcal{E}_n$ (cf. Equation (26)) as target space (left) and SafeOpt sampling only from the set of expanders $\mathcal{G}_n^{\text{SafeOpt}}$ (right).

I.3 Jumping Past Local Barriers

In this additional experiment we demonstrate that ITL is able to extrapolate safety beyond local unsafe "barriers", which is a fundamental limitation of Lipschitz-based methods such as SafeOpt. We consider the ground truth function and prior statistical model shown in Figure 8. Note that initially, there are three disjoint safe "regions" known to the algorithm corresponding to two of the three safe "bumps" of the ground truth function. In this experiment, the main challenge is to "jump past" the local barrier separating the leftmost and initially unknown safe "bump".

Figure 9 shows the sampled points during the first 100 iterations of SafeOpt and ITL. Clearly, SafeOpt does not discover the third safe "bump" while ITL does. Indeed, it is a fundamental limitation of Lipschitz-based methods that they can never "jump past local barriers", even if the oracle Lipschitz constant were to be known and tight (i.e., locally accurate) around the barrier. This is because Lipschitz-based methods expand to the point $\boldsymbol{x}$ based on a reference point $\boldsymbol{x}'$, and by definition, if $\boldsymbol{x}$ is added to the safe set so are all points on the line segment between $\boldsymbol{x}$ and $\boldsymbol{x}'$. Hence, if there is a single point on this line segment which is unsafe (i.e., a "barrier"), the algorithm will *never* expand past it. This limitation does not exist for kernel-based algorithms as expansion occurs in function space.

Moreover, note that for a non-stationary kernel such as in this example, ITL samples the "closest points" in function space rather than Euclidean space. We observe that SafeOpt still samples "locally at the boundary" whereas ITL samples the most informative point which in this case is

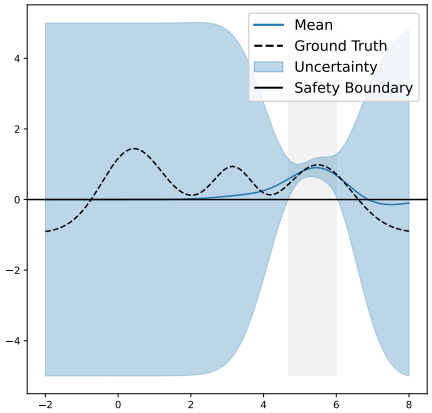

Figure 10: Ground truth and prior well-calibrated model in 1d synthetic experiment. The function serves simultaneously as objective and as constraint. The light gray region denotes the initial safe set $\mathcal{S}_0$.

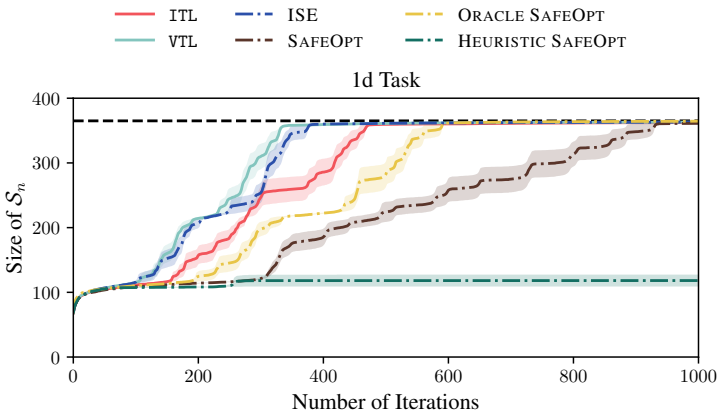

Figure 11: Size of $\mathcal{S}_n$ in 1d synthetic experiment. The dashed black line denotes the size of $\mathcal{S}^\star$. In this task, "discovering" the optimum is closely linked to expansion of the safe set, and Heuristic SafeOpt fails since it does not expand the safe set sufficiently.

the local maximum of the sinusoidal function. In other words, ITL adapts to the geometry of the function. This generally leads us to believe that ITL is more capable to exploit (non-stationary) prior knowledge than distance-based methods such as SafeOpt.

I.4 Experiment Details

I.4.1 Synthetic Experiments

1d task Figure 10 shows the objective and constraint function, as well as the prior. We discretize using 500 points. The main difficulty in this experiment lies in sufficiently expanding the safe set to discover the global maximum. Figure 11 plots the size of the safe set $\mathcal{S}_n$ for the compared algorithms, which in this experiment matches the achieved regret closely.

2d task We model our constraint in the form of a spherical "island" where the goal is to get a good view of the coral reef located to the north-east of the island while staying in the interior of the island during exploration (cf. Figure 12). The precise objective and constraint functions are unknown to the agent. Hence, the agent has to gradually and safely update its belief about boundaries of the "island" and the location of the coral reef. The prior is obtained by a single observation within the center of the island $[-0.5, 0.5]^2$. We discretize using 2 500 points.

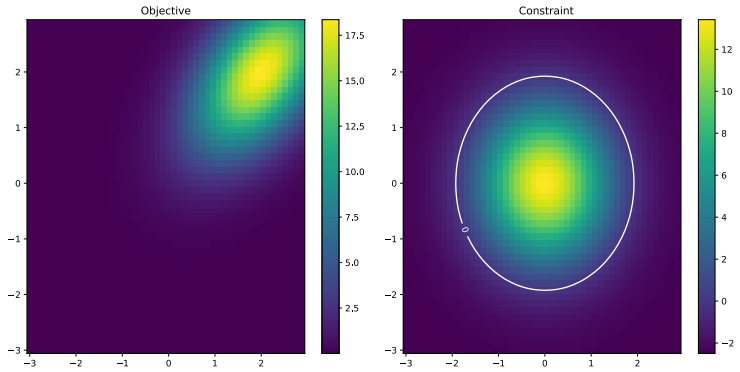

Figure 12: Ground truth in 2d synthetic experiment.

I.4.2 Safe Controller Tuning for Quadcopter

Modeling the real-world dynamics We learn a feedback policy (i.e., "control gains") to compensate for inaccuracies in the initial controller. In our experiment, we model the real world dynamics and the adjusted model using the PD control feedback (Widmer et al., 2023),

$$\boldsymbol{\delta}_t(\boldsymbol{x}) \stackrel{\text{def}}{=} (\boldsymbol{x}^\star - \boldsymbol{x})[(\boldsymbol{s}^\star - \boldsymbol{s}_t)\ (\dot{\boldsymbol{s}}^\star - \dot{\boldsymbol{s}}_t)], \tag{44}$$

where $\boldsymbol{x}^\star$ are the *unknown* ground truth disturbance parameters, and $\boldsymbol{s}^\star$ and $\dot{\boldsymbol{s}}^\star$ are the desired state and state derivative, respectively. This yields the following ground truth dynamics:

$$\boldsymbol{s}_{t+1}(\boldsymbol{x}) = \boldsymbol{T}(\boldsymbol{s}_t, \boldsymbol{u}_t + \boldsymbol{\delta}_t(\boldsymbol{x})). \tag{45}$$

The feedback parameters $\boldsymbol{x} = [\boldsymbol{x}_p\ \boldsymbol{x}_d]^\top$ can be split into $\boldsymbol{x}_p$ tuning the state difference which are called *proportional parameters* and $\boldsymbol{x}_d$ tuning the state derivative difference which are called *derivative parameters*. We use the "critical damping" heuristic to relate the proportional and derivative parameters: $\boldsymbol{x}_d = 2\sqrt{\boldsymbol{x}_p}$. We thus consider the restricted domain $\mathcal{X} = [0, 20]^4$ where each dimension corresponds to the proportional feedback to one of the four rotors.

Ground truth disturbance parameters are sampled from a chi-squared distribution with one degree of freedom (i.e., the square of a standard normal distribution), $\boldsymbol{x}_p^\star \sim \chi_1^2$, and $\boldsymbol{x}_d^\star$ is determined according to the critical damping heuristic.

The learning problem The goal of our learning problem is to move the quadcopter from its initial position $\boldsymbol{s}(0) = [1\ 1\ 1]^\top$ (in Euclidean space with meter as unit) to position $\boldsymbol{s}^\star = [0\ 0\ 2]^\top$. Moreover, we aim to stabilize the quadcopter at the goal position, and therefore regularize the control signal towards an action $\boldsymbol{u}^\star$ which results in hovering (approximately) without any disturbances. We formalize these goals with the following objective function:

$$f^\star(\boldsymbol{x}) \stackrel{\text{def}}{=} -\sigma\left(\sum_{t=0}^{T} \|\boldsymbol{s}^\star - \boldsymbol{s}_t(\boldsymbol{x})\|_{\boldsymbol{Q}}^2 + \|\boldsymbol{u}^\star - \boldsymbol{u}_t(\boldsymbol{x})\|_{\boldsymbol{R}}^2\right) \tag{46}$$

where $\sigma(v) \stackrel{\text{def}}{=} \tanh((v - 100)/100)$ is used to smoothen the objective function and ensure that its range is $[-1, 1]$. The non-smoothed control objective in Equation (46) is known as a *linear-quadratic regulator* (LQR) which we solve exactly for the undisturbed system using ILQR (Tu et al., 2023). Finally, we want to ensure at all times that the quadcopter is at least 0.5 meter above the ground, that is,

$$g^\star(\boldsymbol{x}) \stackrel{\text{def}}{=} \min_{t \in [T]} \boldsymbol{s}_t^z(\boldsymbol{x}) - 0.5 \tag{47}$$

where we denote by $\boldsymbol{s}_t^z$ the z-coordinate of state $\boldsymbol{s}_t$.

We use a time horizon of $T = 3$ seconds which we discretize using 100 steps. The objective is modeled by a zero-mean GP with a Matérn($\nu = 5/2$) kernel with lengthscale 0.1, and the constraint is modeled by a GP with mean -0.5 and a Matérn($\nu = 5/2$) kernel with lengthscale 0.1. The prior is obtained by a single observation of the "safe seed" $[0\ 0\ 0\ 10]^\top$.

Adaptive discretization We discretize the domain $\mathcal{X}$ adaptively using coordinate LINEBO (Kirschner et al., 2019). That is, in each iteration, one of the four control dimensions is selected uniformly at random, and the active learning oracle is executed on the corresponding one-dimensional subspace.

Safety Using the (unsafe) constrained BO algorithm EIC (Gardner et al., 2014) leads constraint violation,[18] while ITL and VTL do not violate the constraints during learning for any of the random seeds.

Hyperparameters The observation noise is Gaussian with standard deviation $\rho = 0.1$. We let $\beta = 10$. The control target is $\boldsymbol{u}^\star = [1.766\ 0\ 0\ 0]^\top$.

The state space is 12-dimensional where the first three states correspond to the velocity of the quadcopter, the next three states correspond to its acceleration, the following three states correspond to its angular velocity, and the last three states correspond to its angular velocity in local frame. The LQR parameters are given by

$$\boldsymbol{Q} = \operatorname{diag}\{1, 1, 1, 1, 1, 1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1\} \quad \text{and}$$
$$\boldsymbol{R} = 0.01 \cdot \operatorname{diag}\{5, 0.8, 0.8, 0.3\}.$$

The quadcopter simulation was adapted from Chandra (2023).

Each one-dimensional subspace is discretized using 2 000 points.

Random seeds We repeat the experiment for 25 different seeds where the randomness is over the ground truth disturbance, observation noise, and the randomness in the algorithm.

[18] On average, 1.6 iterations of the first 50 violate the constraints.

Table 1: Magnitudes of γ_n for common kernels. The magnitudes hold under the assumption that $\mathcal{X}$ is compact. Here, B_ν is the modified Bessel function. We take the magnitudes from Theorem 5 of Srinivas et al. (2009) and Remark 2 of Vakili et al. (2021). The notation $\widetilde{O}(\cdot)$ subsumes log-factors. For $\nu = 1/2$, the Matérn kernel is equivalent to the Laplace kernel. For $\nu \to \infty$, the Matérn kernel is equivalent to the Gaussian kernel. The functions sampled from a Matérn kernel are $\lceil \nu \rceil - 1$ mean square differentiable. The kernel-agnostic bound follows by simple reduction to a linear kernel in $|\mathcal{X}|$ dimensions.

Kernel	$k(\boldsymbol{x}, \boldsymbol{x'})$	γ_n
Linear	$\boldsymbol{x}^\top \boldsymbol{x'}$	$O(d \log(n))$
Gaussian	$\exp\left(-\frac{\left\lVert \boldsymbol{x}-\boldsymbol{x'} \right\rVert_2^2}{2h^2}\right)$	$\widetilde{O}\left(\log^{d+1}(n)\right)$
Laplace	$\exp\left(-\frac{\left\lVert \boldsymbol{x}-\boldsymbol{x'} \right\rVert_1}{h}\right)$	$\widetilde{O}\left(n^{\frac{d}{1+d}} \log^{\frac{1}{1+d}}(n)\right)$
Matérn	$\frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}\left\lVert \boldsymbol{x}-\boldsymbol{x'} \right\rVert_2}{h}\right)^\nu B_\nu \left(\frac{\sqrt{2\nu}\left\lVert \boldsymbol{x}-\boldsymbol{x'} \right\rVert_2}{h}\right)$	$\widetilde{O}\left(n^{\frac{d}{2\nu+d}} \log^{\frac{2\nu}{2\nu+d}}(n)\right)$
any		$O(\lvert\mathcal{X}\rvert \log(n))$