Delft University of Technology

Efficiently Computable Safety Bounds for Gaussian Processes in Active Learning

Tebbe, Jörn; Zimmer, Christoph; Steland, Ansgar; Lange-Hegermann, Markus; Mies, Fabian

Publication date
2024

Document Version
Final published version

Published in
Proceedings of the 27th International Conference on Artificial Intelligence and Statistics (AISTATS) 2024, Valencia, Spain

Citation (APA)
Tebbe, J., Zimmer, C., Steland, A., Lange-Hegermann, M., & Mies, F. (2024). Efficiently Computable Safety Bounds for Gaussian Processes in Active Learning. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics (AISTATS) 2024, Valencia, Spain* (Vol. 238). (Proceedings of Machine Learning Research). PMLR.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Efficiently Computable Safety Bounds for Gaussian Processes in Active Learning

Jörn Tebbe
OWL University of Applied Sciences and Arts

Christoph Zimmer
Bosch Center for Artificial Intelligence

Ansgar Steland
RWTH Aachen University

Markus Lange-Hegermann
OWL University of Applied Sciences and Arts

Fabian Mies
TU Delft

Abstract

Active learning of physical systems must commonly respect practical safety constraints, which restricts the exploration of the design space. Gaussian Processes (GPs) and their calibrated uncertainty estimations are widely used for this purpose. In many technical applications the design space is explored via continuous trajectories, along which the safety needs to be assessed. This is particularly challenging for strict safety requirements in GP methods, as it employs computationally expensive Monte-Carlo sampling of high quantiles. We address these challenges by providing provable safety bounds based on the adaptively sampled median of the supremum of the posterior GP. Our method significantly reduces the number of samples required for estimating high safety probabilities, resulting in faster evaluation without sacrificing accuracy and exploration speed. The effectiveness of our safe active learning approach is demonstrated through extensive simulations and validated using a real-world engine example.

1 INTRODUCTION

Active learning is a machine learning technique that involves selecting the most informative examples from a large unlabeled dataset and requesting their labels from an oracle, e.g., a human annotator or a costly experiment, to improve the performance of a learning algorithm (Settles, 2009; Tharwat and Schenck, 2023). The goal of active learning is to reduce the number of labeled examples needed to achieve high accuracy.

In many engineering scenarios, the next experiment must not only be informative, but also adhere to *practical safety constraints* (Sui et al., 2018; Berkenkamp et al., 2016; Baumann et al., 2021). For example, we consider the control of a high-pressure fluid system for fuel injection in combustion engines, see Section 5.3. The experimental conditions need to be chosen such that a critical pressure threshold is not exceeded. Challengingly, the exact effect of the controls on the pressure in the system is unknown and needs to be learned simultaneously. Safe active learning aims to balancing the trade-off between exploration performance and safety, ensuring that the selected examples are not only informative but can also be obtained securely.

An established approach in safe active learning is to model the unknown functional relation via Gaussian processes (GPs) (Rasmussen et al., 2006). This allows for Bayesian uncertainty quantification and, hence, for informed decisions about where to sample next (Schreiter et al., 2015; Zimmer et al., 2018; Li et al., 2022). This makes GPs a powerful tool for maximizing physical experiments' information while minimizing risk.

A typical additional challenge is exploring *dynamical systems*, where exploration along trajectories instead of single datapoints is necessary. In our engine example a controller continuously adapts engine speed and rail pressure, and similarly the path in robot exploration needs to be safe. In such applications, the entire

Proceedings of the 27$^{\text{th}}$ International Conference on Artificial Intelligence and Statistics (AISTATS) 2024, Valencia, Spain. PMLR: Volume 238. Copyright 2024 by the author(s).

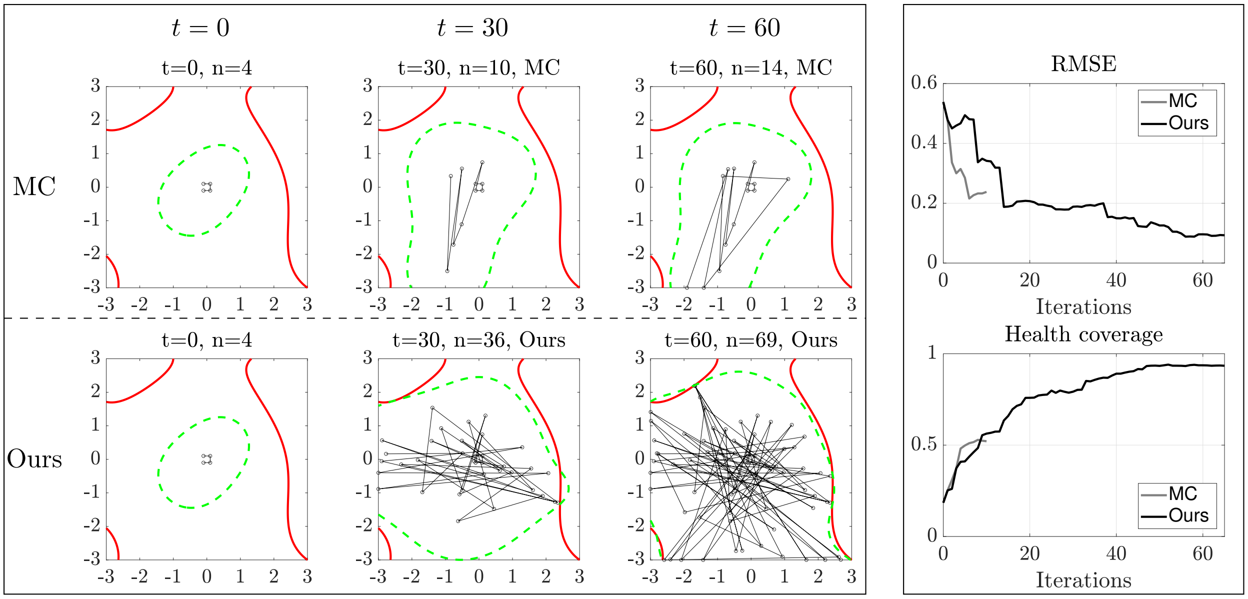


Figure 1: By providing better estimates, we obtain accurate error bounds with much fewer MC samples. This reduction in computation time for the safety evaluation allows more time to obtain measurements for Safe Active Learning. The three left columns present visual representations of a Safe Active Learning task. Each column corresponds to a different algorithm runtime t, along with the respective number of training points n, resulting in $n-1$ explored trajectories. In these plots, the green dashed region marks the area classified as safe by the GP, while the space outside the red boundary indicates the ground truth unsafe region. The rightmost column provides comparisons of two crucial metrics, contingent on the number of iterations. These comparisons underscore the superiority of our approach in enhancing the effectiveness of the Safe Active Learning process by allowing more iteration in the same time.

trajectory needs to be safe.

GPs are again attractive for exploration along trajectories. They naturally induce a univariate GP as posterior of the safety constraint along the trajectory. Now, the safety probability is the probability of this univariate GP being safe which is analytically intractable. The state of the art approach is to *estimate the safety of paths* by generating enough Monte-Carlo (MC) samples from the GP on a finite set of points on the path (Zimmer et al., 2018, 2020).

However, this has a major drawback: obtaining high safety guarantees requires a large number of MC samples, as one needs enough samples in the tails of the distribution. This exposes a trade-off between the safety of the trajectory, the confidence in the safety assessment, and the computational costs. The latter is especially important if decisions need to be taken quickly, as otherwise costly measurement equipment and staff is idly waiting (Sandmeier, 2022; Thewes et al., 2016).

This paper tackles this trade-off by a novel, comparatively tight, and computationally efficient algorithms to compute provable *probabilistic upper bounds* on GP maxima evaluated on a discretization. We develop a variant of the Borell-TIS bound, make it well-suited for practical applications, and adapt this bound to allow for non-centered GPs. The Borell-TIS inequality allows us to draw conclusions about the far tails of the posterior distribution based on the median and the maximal variance, which can be estimated adaptively and reliably with comparatively few MC samples. Our algorithm uses this inequality adaptively, and hence achieves high precision with minimal computational overhead, enabling faster and more efficient learning. This paper makes the following contributions:

- We reduce the *safety assessment for a non-centered GP* to a centered GP, see Remark 3 and Figure 2.
- We propose a computationally efficient method for the safety assessment using an *adaptive MC sampling scheme.*
- We rigorously prove in Section 4 and empirically demonstrate on various examples in Section 5 the *safety guarantees* of our approach.
- We demonstrate that our approach is significantly faster to compute than state-of-the-art safe active learning techniques, hence it enables *more exploration in the same time* (see Section 5).

2 PRELIMINARIES

2.1 Gaussian processes (GPs)

A Gaussian process (GP) $g = \mathcal{GP}(\mu, k)$ is a stochastic process characterized by its mean function $\mu\colon \mathbb{R}^d \to \mathbb{R}$ and covariance function $k\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. Conditioning $g = \mathcal{GP}(\mu, k)$ on a dataset $(x, y) \in \mathbb{R}^{n \times (d+1)}$ yields a posterior GP with mean and covariance functions

$$\mu^*(x^*) = \mu(x^*) + k(x^*, x)(K + \sigma_n^2 I)^{-1} y$$
$$\Sigma(x_1, x_2) = k(x_1, x_2) - k(x_1, x)(K + \sigma_n^2 I)^{-1} k(x, x_2)$$

with covariance matrix $K = (k(x_i, x_j))_{i,j} \in \mathbb{R}^{n \times n}$, and $k(x^*, x) \in \mathbb{R}^{1 \times n}$, $k(x_1, x) \in \mathbb{R}^{1 \times n}$, $k(x, x_2), k(x, x^*) \in \mathbb{R}^{n \times 1}$ and noise variance σ_n^2 (Rasmussen et al., 2006). In practice, GPs are parameterized by hyperparameters $\theta \in \mathbb{R}^p$. These hyperparameters are adapted to data by minimizing the negative log-likelihood $p(y|x, \theta)$. We use the squared exponential covariance function

$$k_{\text{SE}}(x_1, x_2) = \sigma_f^2 \exp\left(-\frac{1}{2}\frac{(x_1 - x_2)^2}{\ell^2}\right)$$

for the GP priors.

Sampling a GP $g = \mathcal{GP}(\mu, k)$ at a finite number of points $x_1^*, \ldots, x_m^* \in \mathbb{R}^d$ amounts to sampling from the m-dimensional Gaussian distribution $\mathcal{N}(v, K^*)$ with $v_i = \mu(x_i^*)$ and $(K^*)_{i,j} = k(x_i^*, x_j^*)$ for $i, j \in \{1, \ldots, m\}$. The computational costs for generating M samples consist of $\mathcal{O}(m^3)$ operations for a preliminary Cholesky decomposition of the covariance matrix K^* and $\mathcal{O}(Mm^2)$ operations to simulate the samples, in addition to the cost of computing the posterior.

2.2 Safe Active Learning

Safe Active Learning selects sample locations $x_1, \ldots, x_n$ that maximize information content, constrained on safety (Schreiter et al., 2015; Zimmer et al., 2018; Li et al., 2022). Entropy is a measure of information that is particularly suited for GPs as the entropy of a new point x^* is in monotonous bijection to its predictive variance $\sigma^2(x^*)$. Therefore, the core of safe active learning is a constrained optimization problem:

$$x_{n+1} = \text{argmax}_{x^* \in \mathcal{X}}\ \sigma(x^*) \quad \text{s.t.} \quad P_{\text{unsafe}}(x^*) \le \alpha$$

where a small $0 < \alpha \le 1$ denotes the maximal desired probability of unsafety and $\mathcal{X} \subseteq \mathbb{R}^d$ is the operation area of the system. We consider the usual case that the safety of a point x^* characterized in terms of a (unknown) safety indicator function $f : \mathbb{R}^d :\to \mathbb{R}$. That is, an operational setting x^* is safe if $z = f(x^*) \ge z_{\min}$. By shifting the mean value accordingly, we may set $z_{\min} = 0$ without loss of generality. We assume the safety indicator to be experimentally measurable, so that we can model it via a posterior GP $\widehat{f} \sim GP(\widehat{\mu}, \widehat{\Sigma})$, where the posterior parameters $\widehat{\mu}$ and $\widehat{\Sigma}$ depend on the previously explored samples $x_1, \ldots, x_n$ and their evaluations $z_i = f(x_i)$, as described in Section 2.1. Then we denote by $P_{\text{unsafe}}(x^*)$ the posterior probability

$$P_{\text{unsafe}}(x^*) = 1 - \int_{z \ge 0} \mathcal{N}(z; \widehat{\mu}(x^*), \widehat{\Sigma}(x^*, x^*)) dz. \quad (1)$$

In case of active learning in *dynamic systems* (Zimmer et al., 2018), the exploration is usually conducted along parameterized trajectories $\tau(t)_{t \in [0,1]}$ instead of points x^*, e.g. the trajectory τ can be a linear ramp leading to some end point of interest. The active learning task then consists in choosing a sequence $\tau_1, \tau_2, \ldots$ of trajectories which maximize information content, constrained by the safety requirement $P_{\text{unsafe}} \le \alpha$. A measurement is conducted at the endpoint of the trajectory τ, and information is then measured with the posterior GPs $\widehat{f}$ predictive variance $\widehat{\sigma}(\tau(1))$. The considered probability of the trajectory being unsafe is

$$P_{\text{unsafe}}(\tau) = P\left(\inf_{t \in [0,1]} Z_t \le 0\right), \quad (2)$$

where Z_t is a sample of the posterior GP.

In the same framework, it is also possible to conduct measurements along the trajectory at locations $\tau(t_1), \ldots, \tau(t_m)$ for $t_1, \ldots, t_m \in [0, 1]$. The information of these measurements may be expressed in terms of the predictive covariance matrix $\widehat{\Sigma}(\tau(t_1), \ldots, \tau(t_m))$, and quantified via its trace or determinant.

3 RELATED WORK

Estimating bounds for a GP is a crucial task for exploration in safety critical environments. Other approaches in the literature consider a bound of the RKHS norm of the safety function in order to create confidence intervals. While Sui et al. (2018) propose to additionally use estimated Lipschitz-constants of the safety function, Bottero et al. (2022) overcomes this practically strong assumption. In Lederer et al. (2019), the authors provide uniform error bounds based on Lipschitz constants estimations. Schreiter et al. (2015) proposes a method to obtain pointwise safety in safe active learning. All these works consider pointwise safety instead of safety on continuous trajectories, Zimmer et al. (2018) extend safety consideration to trajectories. This is improved in Zimmer et al. (2020) by an adaptive discretization scheme. Our novel approaches are compatible with both these papers and for simplicity we compare ourselves to the former one.

Moreover, Cardelli et al. (2019) consider safety for compact sets to detect adversarial attacks on the data.

For this purpose they use a variant of the Borell-TIS inequality (Adler and Taylor, 2007) which bounds the mean of the supremum of a GP using Dudley's theorem (Dudley, 1967). We show in Section 5.1, that these bounds are inferior, compared to our proposed method.

While our experiments use standard GPs with squared exponential covariance function, our methods directly extend to usual variants and approximations to GPs. This includes any separable covariance function, including specific ones constructed from kernel search (Duvenaud et al., 2013; Bitzer et al., 2022), geometry (Borovitskiy et al., 2020), differential equations (Besginow and Lange-Hegermann, 2022; Härkönen et al., 2023), for high dimensional modeling (Duvenaud et al., 2011), kernels building on Fourier frequencies (Lázaro-Gredilla et al., 2010), or symmetry (Holderrieth et al., 2021). Furthermore, this includes GPs for big data, be it via variational approximations (Titsias, 2009; Hensman et al., 2013, 2017), via kernel approximations (Wilson and Nickisch, 2015), or improved linear algebra (Gardner et al., 2018; Wang et al., 2019).

The safety assessment in (1) reduces to a Gaussian integral under linear constraints which has been considered by Genz (1992). This method has been extended to high dimensional integrals ($m > 100$) of mainly small areas (Botev, 2017; Gessner et al., 2020). These methods use variants of Monte-Carlo sampling, but prior experiments indicate that they have inferior performance as our proposed methods.

Adaptive stopping Monte-Carlo schemes have been used in the literature for estimating statistical quantities (Mnih et al., 2008). Our stopping scheme extends via exploiting an additional binomial structure.

4 UNIFORM TAIL BOUNDS FOR GAUSSIAN PROCESSES

When exploring the experimental space along a trajectory $(\tau(t))_{t\in[0,1]} \in \mathbb{R}^d$, we want to be reasonably certain that a safety indicator $f(\tau(t))$ does not fall below a critical threshold $z_{\min} = 0$. Denote the trajectories of the posterior distribution for $f(\tau(t))_{t\in[0,1]}$ by $(Z_t)_{t\in[0,1]}$, i.e. Z_t is a GP with one-dimensional input, mean function $\mathbb{E}(Z_t) = \mu_t \in \mathbb{R}, t \in [0,1]$, and covariance function $\mathrm{Cov}(Z_s, Z_t) = \Sigma_{s,t}$, derived according to the formulas in Section 2.1. Hence, the (posterior) probability P_{unsafe} of the trajectory τ being unsafe is given by (2). For computational purposes, the continuous trajectory $\tau(t)$ needs to be discretized by considering only a subset $T \subset [0,1]$, such that

$$P_{\text{unsafe}}(\tau) \approx P\left(\inf_{t\in T} Z_t \le 0\right) =: P^*(\tau).$$

Typically, $T = \{t_1, \ldots, t_m\}$ is a finite set for computational purposes. However, we want to emphasize that our adaptive sampling scheme and the analytical bounds presented in Section 4.2 also hold for the continuous case $T = [0,1]$.

For safe exploration, we want to accept the trajectory $\tau(t)_{t\in T}$ as being safe only if $P^*(\tau) \le \alpha$, for some small $\alpha \in (0,1]$. This section describes reliable upper bounds on P^* which should be (i) fast to evaluate computationally and (ii) as sharp as possible.

4.1 Adaptive Monte-Carlo Sampling (AMC)

The state of the art chooses a potentially non-equidistant discretization of $[0,1]$ given by $0 \le t_1 < \ldots < t_m \le 1$ and simulate a potentially large number, M, of trajectories $Z_{t,1}, \ldots, Z_{t,M}$ evaluated only on the discretization $T = \{t_1, \ldots, t_m\}$. The computational cost is dominated by $\mathcal{O}(Mm^2)$ for large enough M, see Section 2.1. We estimate the probability of unsafe trajectories by their proportion in the sampled trajectories

$$\widehat{P}_{\text{MC}}(\tau, M) = \frac{1}{M}\sum_{i=1}^{M} \mathbf{1}\left(\min_{j=1,\ldots,m} Z_{t_j,i} \le 0\right).$$

How many MC samples M are necessary? Since $\mathrm{Var}(\widehat{P}_{\text{MC}}) \le P^*/M$, the relative error compared to the safety threshold is of the order

$$|\widehat{P}_{\text{MC}} - P_{\text{MC}}|/\alpha = \mathcal{O}(\sqrt{P^*}/\sqrt{M\,\alpha}).$$

Thus, for trajectories which are barely safe, $P^* \approx \alpha$, we should perform $M \asymp \alpha^{-1}$ MC iterations. That is, for strict safety requirements $\alpha \approx 0$, the MC approach is computationally expensive. Determining the exact number of required MC samples is non-trivial: If $P^* \ll \alpha$, i.e. if the trajectory is very safe, then few samples suffice as the variance of $\widehat{P}_{\text{MC}}$ is small. If, on the other hand, $P^* \gg \alpha$, it is also sufficient to draw rather few samples, as the mean of $\widehat{P}_{\text{MC}}$ is far away from the decision boundary. Specifically in the critical regime $P^* = \alpha + O(\delta)$ for some small $|\delta|$ it is hard to decide whether the trajectory τ is indeed safe or unsafe; the smaller $|\delta|$, the more samples are needed. Unfortunately, given a candidate trajectory τ, we do not know P^* in advance.

We suggest to determine the sample size adaptively: generate the MC samples sequentially and perform an online test for the hypothesis $H_0 : P^*(\tau) \ge \alpha$, and to stop sampling as soon as H_0 is rejected, classifying the trajectory as safe. At the same time, we stop when $H_0' : P^*(\tau) \le \alpha$ is rejected, classifying the trajectory as unsafe. We suggest to sequentially increase the sample size $M_1 < M_2 < \ldots$, and to stop sampling at step r^* with sample size M_{r^*} for

$$\begin{aligned} r^* = \inf\Big\{r\ &: \widehat{P}^+_{\mathrm{MC}}(\tau, M_r, r, \epsilon, \alpha) < \alpha \text{ or} \\ &\widehat{P}^-_{\mathrm{MC}}(\tau, M_r, r, \epsilon, \alpha) > \alpha)\Big\}, \\ \widehat{P}^+_{\mathrm{MC}}(\tau, M_r, r, \epsilon, \alpha) &:= \widehat{P}_{\mathrm{MC}}(\tau, M_r) + \sqrt{\alpha(1-\alpha)}c_r, \\ \widehat{P}^-_{\mathrm{MC}}(\tau, M_r, r, \epsilon, \alpha) &:= \widehat{P}_{\mathrm{MC}}(\tau, M_r) - \frac{c_r^2}{4} - c_r\sqrt{\alpha}, \\ \text{with } c_r &= \sqrt{\frac{2}{M_r}\left|\log\frac{6\epsilon}{\pi^2 r^2}\right|} \end{aligned}$$

where $\epsilon > 0$ should be small. Stopping due to $\widehat{P}^+_{\mathrm{MC}}$ resp. $\widehat{P}^-_{\mathrm{MC}}$ classifies τ as safe resp. unsafe. Indeed, this procedure can control the probability of falsely classifying an unsafe trajectory as safe and vice versa.

Theorem 1. *Let Q denote the probability w.r.t. the MC sampling and let $\epsilon \in (0,1)$. If $P^*(\tau) \geq \alpha$ then*

$$Q\left(\exists r \in \mathbb{N} : \widehat{P}^+_{\mathrm{MC}}(\tau, M_r, r, \epsilon, \alpha) < \alpha\right) \leq \epsilon$$

and if $P^*(\tau) \leq \alpha$, then

$$Q\left(\exists r \in \mathbb{N} : \widehat{P}^-_{\mathrm{MC}}(\tau, M_r, r, \epsilon, \alpha) > \alpha\right) \leq \epsilon.$$

We call this method Adaptive Monte-Carlo (AMC). Since $c_r \to 0$ and $\widehat{P}_{\mathrm{MC}}(\tau, M_r) \to P^*$, the stopping time r^* is almost surely finite if $P^* \neq \alpha$. Theorem 1 guarantees the above method decides correctly with probability $1-\epsilon$ if $P^* \neq \alpha$; we may choose ϵ small. On the other hand, Theorem 1 also implies that $P(r^* = \infty) \geq 1 - \epsilon$ for the edge case $P^* = \alpha$. Thus, in practice, we impose an upper bound on r^* and classify a trajectory as unsafe if this bound is exceeded.

4.2 Analytical bound

We suggest to bound P^* via the Borell-TIS inequality for GPs. It asserts the remarkable result that the supremum of a GP shifted by the mean or median of suprema of samples has subgaussian tails.

Theorem 2 (Borell-TIS inequality)**.** *Let $X_t, t \in T$, be a centered separable GP with index set T, and maximal pointwise variance $\sigma^2 = \sup_{t\in T} Var(X_t)$. Denote $m(X) = \mathrm{median}\left(\sup_{t\in T} X_t\right)$, $\mu(X) = \mathbb{E}\left(\sup_{t\in T} X_t\right)$ and Φ as the standard Gaussian distribution function. Then, for any $u \geq 0$,*

$$P\left[\sup_{t\in T} X_t > u + m(X)\right] \leq \left[1 - \Phi(u/\sigma)\right] \qquad \text{(B.1)}$$

$$\leq \tfrac{1}{2}\exp(-\tfrac{1}{2}u^2/\sigma^2), \qquad \text{(B.2)}$$

$$P\left[\sup_{t\in T} X_t > u + \mu(X)\right] \leq \exp(-\tfrac{1}{2}u^2/\sigma^2). \qquad \text{(B.3)}$$

Typically, (B.1) is the sharpest of the bounds of Theorem 2; cf. Figure 3 for a numerical comparison. See van der Vaart and Wellner (1996) for a proof of Theorem 2. Note that the sharpest inequality (B.1) is derived in the proof of Lemma A.2.2 therein.

Remark 3. In our setting, Theorem 2 is not directly applicable because $Z = \mathcal{GP}(\mu, \Sigma)$ is usually conditioned on previously obtained data and hence of variable mean. We remedy this technical problem as follows: suppose that $\mu_t > 0$ for all $t \in T$, since otherwise $P^* \geq \frac{1}{2}$ which is a too high risk of failure. Then, we may rewrite

$$\inf_{t\in T} Z_t \geq 0 \iff X_t := \frac{Z_t - \mu_t}{-\mu_t} \leq 1 \quad \forall t \in T.$$

Figure 2 exemplifies this centering transformation. Considering $X = \frac{\mu - Z}{\mu} = \mathcal{GP}\left(0, (s,t) \mapsto \frac{\Sigma(s,t)}{\mu(s)\mu(t)}\right)$, it suffices to derive upper tail bounds for

$$P^* = P\left(\inf_{t\in T} Z_t \leq 0\right) = P\left(\sup_{t\in T} X_t \geq 1\right).$$

This discussion, together with Theorem 2 proves the following analytical bound on P^*.

Theorem 4. *Let $Z_t, t \in T$, be a separable GP with index set T and mean function $\mu_t > 0$. Then, if $\widetilde{m} =$* $\mathrm{median}\left(\sup_{t\in T}(\mu_t - Z_t)/\mu_t\right) \leq 1$, *we have*

$$P^*(\tau) \leq P^\dagger(\tau) = 1 - \Phi\left(\tfrac{1-\widetilde{m}}{\widetilde{\sigma}}\right)$$

for $\widetilde{\sigma}^2 = \sup_{t\in T} Var(Z_t)/\mu_t^2$.

4.3 Semi-analytical bound (AB)

In Theorem 4, the median $\widetilde{m}$ and the maximal variance $\widetilde{\sigma}^2$ cannot be computed in closed form. We approximate them via MC sampling along a discretization of Z_t resp. X_t. Thereby, we replace MC sampling of events in tails of probability distributions by much easier MC sampling of a median of the same distribution. This is possible because the Gaussianity of the process allows us to extrapolate to the tail of the distribution using data from the center of its mass, via Theorem 4. As a consequence, we can drastically decrease the number of MC samples required for reliable tail bounds.

Again, we suggest to determine the MC sample size adaptively. To this end, we make use of exact finite sample confidence intervals for the median as follows. Based on $M \in \mathbb{N}$ samples, denote the simulated maxima by $S_i = \max_j X_{t_j,i}$ and let $q_{\beta,M} = q_{\beta,M}(S_1, \ldots, S_M)$ the empirical β-quantile, i.e. the $\lfloor M\beta \rfloor$-th order statistic. For any M, r, and $\epsilon > 0$, and a confidence level $\chi = \chi(r,\epsilon) = 1 - \frac{6\epsilon}{\pi^2 r^2}$, there exist $\beta_\pm = \beta_\pm(M, r, \epsilon)$, with $\beta_- \leq \beta_+$, such that

$$P\left(\widetilde{m} \in [q_{\beta_-,M_r}, \infty)\right) \geq \chi,$$
$$P\left(\widetilde{m} \in (-\infty, q_{\beta_+,M_r}]\right) \geq \chi.$$

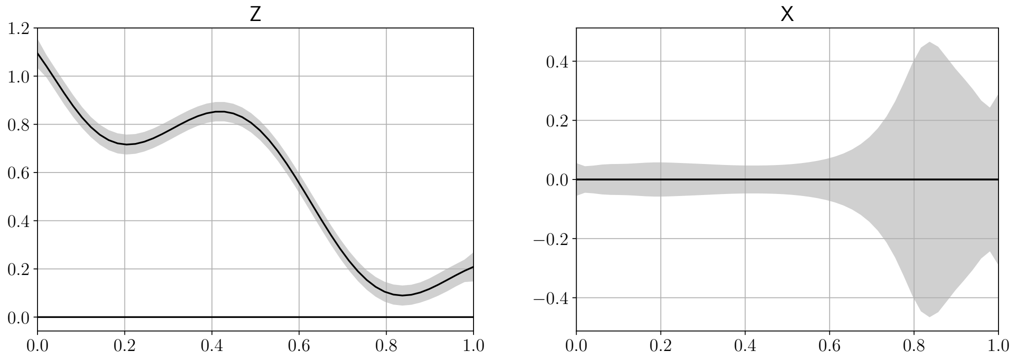


Figure 2: The left diagram illustrates the GP Z_t from Subsection 5.1 via its mean function and pointwise two sigma bands; the right diagram shows the corresponding centered GP X_t resulting from Remark 3. The safety-relevant information of the mean of the original GP Z moves to the covariance of the centered GP X. In particular, the variance of the centered GP X rises where the mean of Z approaches the safety bound zero.

We can use the asymptotic approximation $\beta_\pm \approx \frac{1}{2} \pm \Phi^{-1}(\chi)/\sqrt{4\,M}$ (Conover, 1999, p. 144). Thus, as in Section 4.1, we suggest to sequentially increase the sample size $M_1 < M_2 < \ldots$, and to stop sampling at step r' with sample size $M_{r'}$ for

$$r' = \inf\Big\{r \;:\; \widehat{P}_-^\dagger(M_r, r, \epsilon) > \alpha \text{ or } \widehat{P}_+^\dagger(M_r, r, \epsilon) < \alpha\Big\},$$
$$\text{where} \qquad \widehat{P}_\pm^\dagger(M_r, r, \epsilon) := 1 - \Phi\left(\frac{1 - q_{\beta_\pm, M_r}}{\sigma_m}\right).$$

Theorem 5. *If $P^\dagger(\tau) \geq \alpha$, then for any $\epsilon \in (0,1)$*

$$Q\left(\exists r \in \mathbb{N} : \widehat{P}_+^\dagger(M_r, r, \epsilon) < \alpha\right) \leq \epsilon$$

and if $P^\dagger(\tau) \leq \alpha$, then

$$Q\left(\exists r \in \mathbb{N} : \widehat{P}_-^\dagger(M_r, r, \epsilon) > \alpha\right) \leq \epsilon.$$

where Q denotes probability w.r.t. the MC sampling.

We call this method Adaptive Borell (AB). A general adaptive algorithm is provided in Algorithm 1, while method specific versions are in the supplement.

Algorithm 1 Adaptive Safety evaluation

```
Require: Safety threshold: $\alpha > 0$,
  Threshold for confidence intervals: $\epsilon > 0$,
  Discretization: $t_1, \ldots, t_m$,
  Sample sizes: $0 = M_0 < M_1 < \ldots < M_R$,
  Posterior GP: $X_t$.
  $\widehat{P} \leftarrow 0$
  for $r = 1, \ldots, R$ do
      for $i = M_{r-1} + 1, \ldots, M_r$ do
          Simulate $(X_{t_j,i})_{j=1,\ldots,m}$
          $S_i \leftarrow \max_{j=1,\ldots,m} X_{t_j,i}$
      Update $\widehat{P}$ given $S_i$
      Compute confidence intervals $[\widehat{P}_-, \widehat{P}_+]$
      if $\widehat{P}_+ \leq \alpha$ then
          return SAFE
      else if $\widehat{P}_- \geq \alpha$ then
          return UNSAFE
  return UNSAFE
```

4.4 Hybrid adaptive sampling (ABM)

If the assessed trajectory is either clearly safe ($P^* \ll \alpha$) or clearly unsafe ($P^* \gg \alpha$), the adaptive MC scheme might terminate earlier than the semi-analytical procedure. As both methods use the same samples, there is minimal computational overhead to run both schemes in parallel, with remaining uncertainty $\epsilon/2$ instead of ϵ, and stop as soon as one of them reaches a decision.

Moreover, AB provides an upper bound on the unsafeness probability, $P^\dagger \geq P^*$. That is, even if $P^\dagger > \alpha$ (AB bound classifies a trajectory as unsafe), we might still have $\alpha \geq P^*$ (trajectory is indeed safe). Thus, the AB method is overly confident, which maintains safety, but hinders exploration. As a further improvement, we suggest to run both adaptive methods in parallel, and use the Borell-TIS bound only to conclude safety, but not unsafety. More precisely, we stop sampling at step

$$\begin{aligned} r^\diamond = \inf\Big\{r \;:\; & \;\widehat{P}_+^\dagger(M_r, r, \tfrac{\epsilon}{2}) < \alpha \\ & \text{or } \widehat{P}_{\mathrm{MC}}^+(\tau, M_r, \tfrac{\epsilon}{2}, \alpha) < \alpha \\ & \text{or } \widehat{P}_{\mathrm{MC}}^-(\tau, M_r, \epsilon, \alpha) > \alpha)\Big\}. \end{aligned}$$

In the notation of Algorithm 1, we have a sequential

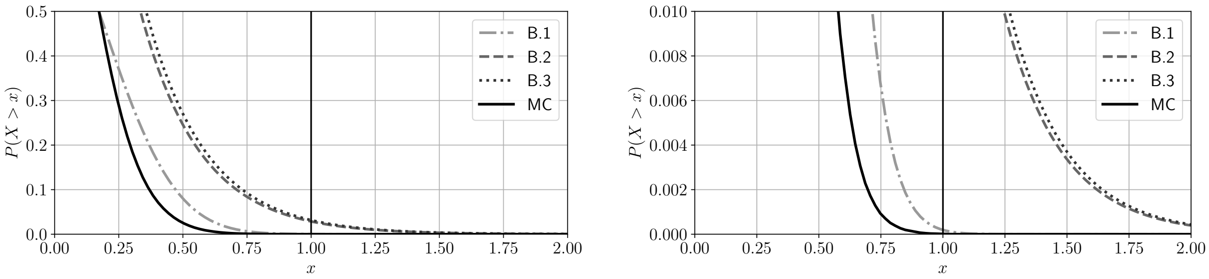


Figure 3: The diagram considers the toy GP from Section 5.1, see also Figure 2 for a visualization of this GP. It shows the tail distribution (complementary cumulative distribution function) $P(X > x)$ for different values of x of different estimations for the supremum of the centered GP. Note, that the safety condition here is $P(X < 1)$ The MC bound can be seen as optimal estimation of $P(X > x)$. Amongst the upper bounds to $P(X > x)$, we see, that the strong Borell inequality B.1 is the sharpest one.

confidence interval given by the bounds

$$\widehat{P}_{+}^{\diamond}(r,\epsilon,\alpha) := \min\left[\widehat{P}_{+}^{\dagger}(M_r, r, \tfrac{\epsilon}{2}),\ \widehat{P}_{\mathrm{MC}}^{+}(\tau, M_r, \tfrac{\epsilon}{2}, \alpha)\right],$$
$$\widehat{P}_{-}^{\diamond}(r,\epsilon,\alpha) := \widehat{P}_{\mathrm{MC}}^{-}(\tau, M_r, \epsilon, \alpha).$$

We call this method Adaptive Borell-Monte-Carlo (ABM). Theorem 1 and Theorem 5 directly imply that the proposed hybrid scheme makes no wrong decisions about safety of a trajectory, with probability $1-\epsilon$, which proves the following corollary.

Corollary 6. *Let Q denote the probability w.r.t. the MC sampling and let $\epsilon \in (0,1)$. If $P^*(\tau) \geq \alpha$ then*

$$Q\left(\exists r \in \mathbb{N} :\ \widehat{P}_{+}^{\diamond}(r,\epsilon,\alpha) < \alpha\right) \leq \epsilon$$

and if $P^*(\tau) \leq \alpha$, then

$$Q\left(\exists r \in \mathbb{N} :\ \widehat{P}_{-}^{\diamond}(r,\epsilon,\alpha) > \alpha\right) \leq \epsilon.$$

5 EXAMPLES

In this section, we test our method using simulated experiments. For multidimensional examples we use linear ramps with equidistant points as trajectories, similar to Zimmer et al. (2018). In our experiments, the computational time budget for each method is fixed, and the number of measurements n_{SAL} which can be taken during one run should be as high as possible. Moreover, we consider different performance metrics to compare the classic Monte-Carlo (MC) method to our novel adaptive Monte-Carlo (AMC) method, adaptive Borell (AB) method, and adaptive Borell-Monte-Carlo (ABM) method.

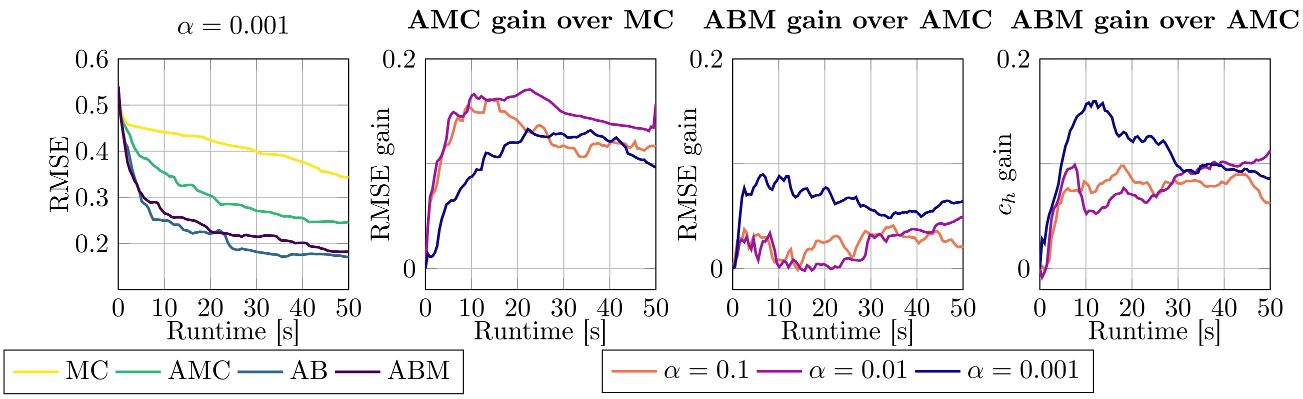
The root-mean-squared error (RMSE) between the obtained GP model and the ground truth indicates the quality of actively learning the behaviour of the real world system. As a second performance metric, we consider the health coverage, defined as the accuracy of a binary classifier which uses the posterior mean $\mu(x)$ to classify the domain as safe or unsafe, i.e.

$$c_h = \frac{\int_{\mathcal{X}} \mathbb{1}_{\mu(x)\geq 0}\mathbb{1}_{z\geq 0}dx + \int_{\mathcal{X}} \mathbb{1}_{\mu(x)<0}\mathbb{1}_{z<0}dx}{\int_{\mathcal{X}} 1dx}.$$

For computational tractability, the integrals are approximated via a fixed number of discrete evaluations.

5.1 Univariate Toy Example

First, we omit the active learning and only assess the quality of our safety evaluation. To this end, we consider the toy example $f\colon [0,1] \to \mathbb{R}, x \mapsto -0.2\sin(10x) - x + 1.1$ with a GP generated by the squared exponential kernel and hyperparameters $\sigma_f = 1$, $\ell^2 = 32^{-1}$ and $\sigma_N^2 = 10^{-3}$. The training points $0 = x_1 \leq x_2 \leq \ldots \leq x_{21} = 1$ are equally spaced, and we consider the safety of the posterior trajectory on the domain $[0,1]$. For a visualization of this GP, see Figure 2. We compare the bounds from our methods with what we regard as the true tail probability: the bound generated by MC sampling with a high number of samples ($M = 10^6$). We discretize the posterior trajectories by 50 equidistant points, i.e. we approximate $P_{\text{unsafe}} \approx P^*$ with $T = \{0, \frac{1}{49}, \ldots, 1\}$ The comparison is depicted in Figure 3. The bound B.1 using the median performs best and is really close to the MC simulation. The bounds given by B.2 and B.3 behave poorly in the tails, as the subgaussian tails are fatter than the Gaussian tails. These numerical results indicate that the bound of AB derived from B.1 is rather tight.



(a) (b) (c) (d)

Figure 4: Consider the Himmelblau's function exploration from Section 5.2. (a) All our three novel adaptive methods (AMC, AB, ABM) improve upon the current state of the art MC with a sufficient sample size (see supplement for exact numbers), when considering the RMSE for high safety requirements $\alpha = 0.001$. (b) Our method AMC improves over MC consistently in RMSE for three different safety requirements. Our favoured method ABM improves further over AMC both for (c) RMSE and (d) detecting the safe region correctly via the health coverage c_h. Results are averaged over 10 independent seeds.

5.2 Himmelblau's function exploration

As an example of an active learning task, we consider exploration of a version of Himmelblau's function $f(x, y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2$ (Himmelblau et al., 2018). In particular, we want to actively learn the function f in the region $[-3, 3]^2$, with the safety constraint $f(x, y) \geq 50$, for a visualization, see Figure 1. Thus, we have a connected safe area with the unsafe area only at the boundary of the given square. For further details on the GP setting, see the supplementary.

The results are shown in Figure 4. The experiments support the theoretical claims, that adaptive methods perform better with rising safety requirements, see Figure 4 (a). This is based on a fewer runtime per iteration due to less MC samples needed to make a provably right decision with high possibility. Indeed, the methods containing B.1 use to make up to 6.5 times as many iterations in the same time, see Figure 1.

5.3 Application: Engine control

We consider a dynamic high-pressure fluid system for fuel injection in combustion engines introduced in Zimmer et al. (2018) as a real world example, see Figure 5. Actuation v_k and the engines' speed n_k are inputs every time step k, and we aim to learn a surrogate model for the rail pressure ψ_k. We inherit the assumption of a nonlinear autoregressive structure $\psi_k = \psi(x_k)$ for $x_k = (n_k, n_{k-1}, n_{k-2}, n_{k-3}, v_k, v_{k-1}, v_{k-3})$. The trajectories linearly interpolate two points equidistantly. For more details about the GP settings, see supplementary. The results are depicted in Figure 6, in analogy to Figure 4. Again, our adaptive methods show improved performance over standard MC, see Figure 6 (a). Especially for higher safety requirements, the methods containing B.1 outperform the others, see Figure 6 (a),(c) and (d). With fewer needed MC samples, AB and ABM can perform more iterations in the same runtime, see the supplementary for details.

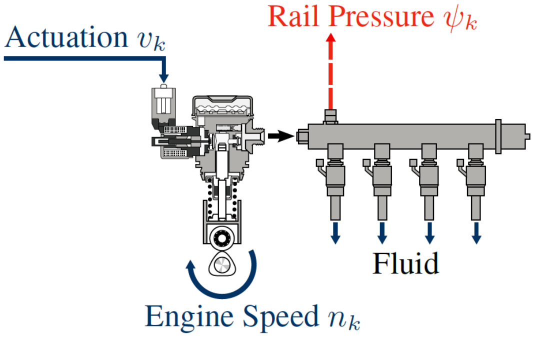


Figure 5: High-pressure fluid injection system with controllable inputs v_k, n_k and measured output ψ_k (picture taken from Zimmer et al. (2018); Tietze et al. (2014))

6 CONCLUSION

In this paper, we explore the derivation of upper bounds on the probabilities of sampled Gaussian processes exceeding prescribed thresholds. Leveraging adaptive techniques, we achieve significant computational efficiency enhancements compared to state-of-the-art Monte-Carlo sampling methods. Furthermore, we ex-

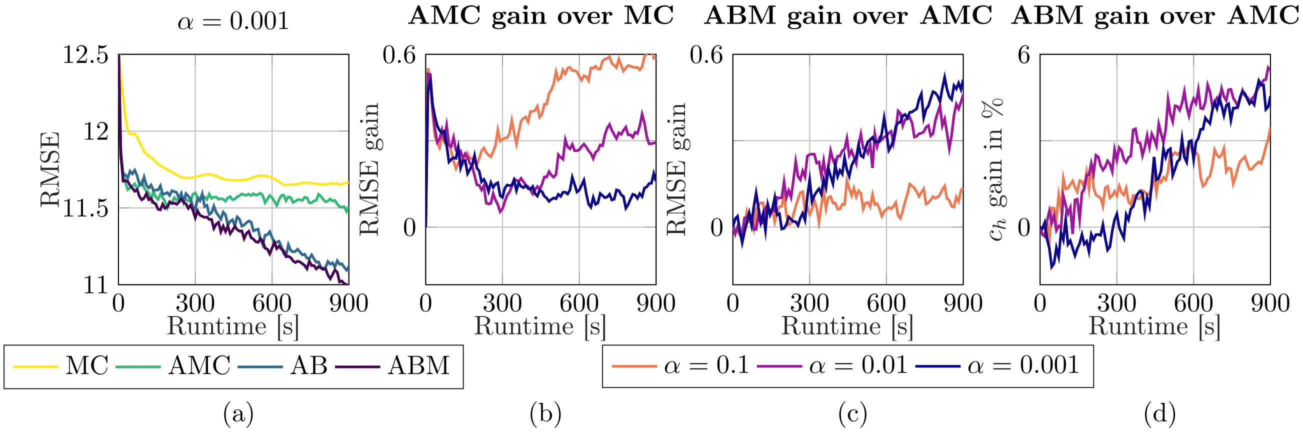


Figure 6: Consider the engine control exploration from Section 5.3. (a) Both our methods AB and ABM using the Borell-TIS inequality B.1 improve upon the current state of the art MC and its adaptive version AMC for high safety requirements $\alpha = 0.001$. (b) The adaptive version AMC improves over MC in RMSE in particular for the lax safety requirement $\alpha = 0.1$. (c) For higher safety requirements $\alpha = 0.01$ and $\alpha = 0.001$ our favoured method ABM improves over AMC and hence also over MC in RMSE. (d) The enhanced RMSE results observed in (c) can be primarily attributed to the improved health coverage c_h. Results are averaged over 10 independent seeds.

tend these advancements by incorporating a variant of the Borell-TIS inequality in conjunction with classical Monte-Carlo sampling. While the Borell-TIS inequality itself entails sampling, it serves to estimate the median rather than the direct estimation of tail probabilities. To facilitate the application of the Borell-TIS inequality, we introduce a centering transformation for Gaussian processes and offer an insightful interpretation. We rigorously establish error bounds for all of our probabilistic methods, ensuring their reliability in practical applications.

Our primary motivation revolves around the domain of safe active learning in dynamic systems. Although our Gaussian process bounds tend to be conservative, potentially resulting in slower exploration, the remarkable reduction in computation time allows for the acquisition of more data points when employing our methods. This advantageous trade-off effectively offsets the conservative nature of our bounds, as empirically demonstrated in our illustrative examples. Notably, our approach proves particularly valuable in scenarios with stringent safety requirements, such as when the trajectory safety probability must exceed 99.9%.

Acknowledgements

Jörn Tebbe is supported by the SAIL project which is funded by the Ministry of Culture and Science of the State of North Rhine-Westphalia under the grant no NW21-059C.
Fabian Mies was employed at RWTH Aachen University while this research was conducted.

References

Adler, R. J. and Taylor, J. E. (2007). *Random Fields and Geometry*. Springer New York.

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

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

Besginow, A. and Lange-Hegermann, M. (2022). Constraining Gaussian processes to systems of linear ordinary differential equations. *NeurIPS*.

Bitzer, M., Meister, M., and Zimmer, C. (2022). Structural kernel search via bayesian optimization and symbolical optimal transport. *NeurIPS*.

Borovitskiy, V., Terenin, A., Mostowsky, P., et al. (2020). Matérn Gaussian processes on riemannian manifolds. *NeurIPS*.

Botev, Z. I. (2017). The normal law under linear restrictions: simulation and estimation via minimax tilting. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 79(1):125–148.

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

Cardelli, L., Kwiatkowska, M., Laurenti, L., and Patane, A. (2019). Robustness guarantees for bayesian inference with Gaussian processes. *AAAI*.

Conover, W. J. (1999). *Practical Nonparametric Statis-*

tics. Probability and Statistics: Applied Probability and Statistics Section. Wiley.

Dudley, R. M. (1967). The sizes of compact subsets of hilbert space and continuity of gaussian processes. *Journal of Functional Analysis*, 1(3):290–330.

Duvenaud, D., Lloyd, J., Grosse, R., Tenenbaum, J., and Zoubin, G. (2013). Structure discovery in nonparametric regression through compositional kernel search. *ICML*.

Duvenaud, D. K., Nickisch, H., and Rasmussen, C. (2011). Additive Gaussian processes. *NeurIPS*.

Gardner, J., Pleiss, G., Weinberger, K. Q., Bindel, D., and Wilson, A. G. (2018). GPytorch: Blackbox matrix-matrix Gaussian process inference with GPU acceleration. *NeurIPS*.

Genz, A. (1992). Numerical computation of multivariate normal probabilities. *Journal of computational and graphical statistics*, 1(2):141–149.

Gessner, A., Kanjilal, O., and Hennig, P. (2020). Integrals over gaussians under linear domain constraints. In *International conference on artificial intelligence and statistics*, pages 2764–2774. PMLR.

Härkönen, M., Lange-Hegermann, M., and Raiţă, B. (2023). Gaussian process priors for systems of linear partial differential equations with constant coefficients. *ICML*.

Hensman, J., Durrande, N., Solin, A., et al. (2017). Variational Fourier features for Gaussian processes. *JMLR*.

Hensman, J., Fusi, N., and Lawrence, N. D. (2013). Gaussian processes for big data. *UAI*.

Himmelblau, D. M. et al. (2018). *Applied nonlinear programming*. McGraw-Hill.

Holderrieth, P., Hutchinson, M. J., and Teh, Y. W. (2021). Equivariant learning of stochastic fields: Gaussian processes and steerable conditional neural processes. *ICML*.

Lázaro-Gredilla, M., Quiñnero-Candela, J., Rasmussen, C. E., and Figueiras-Vidal, A. R. (2010). Sparse spectrum Gaussian process regression. *JMLR*.

Lederer, A., Umlauft, J., and Hirche, S. (2019). Uniform Error and Posterior Variance Bounds for Gaussian Process Regression with Application to Safe Control. *NeurIPS*.

Li, C.-Y., Rakitsch, B., and Zimmer, C. (2022). Safe active learning for multi-output gaussian processes. *AISTATS*.

Mnih, V., Szepesvári, C., and Audibert, J.-Y. (2008). Empirical bernstein stopping. In *Proceedings of the 25th international conference on Machine learning*, pages 672–679.

Rasmussen, C. E., Williams, C. K., et al. (2006). *Gaussian processes for machine learning*. MIT Press.

Sandmeier, N. (2022). *Optimization of adaptive test design methods for the determination of steady-state data-driven models in terms of combustion engine calibration*. Universitätsverlag der Technischen Universität Berlin.

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

Settles, B. (2009). Active learning literature survey. *Computer Sciences Technical Report 1648*.

Sui, Y., Zhuang, V., Burdick, J., and Yue, Y. (2018). Stagewise safe bayesian optimization with gaussian processes. *ICML*.

Tharwat, A. and Schenck, W. (2023). A survey on active learning: State-of-the-art, practical challenges and research directions. *Mathematics*, 11(4):820.

Thewes, S., Krause, M., Reuber, C., Lange-Hegermann, M., Dziadek, R., and Rebbert, M. (2016). Efficient in-vehicle calibration by the usage of automation and enhanced online doe approaches. *Simulation and Testing for Vehicle Technology: 7th Conference*.

Tietze, N., Konigorski, U., Fleck, C., and Nguyen-Tuong, D. (2014). Model-based calibration of engine controller using automated transient design of experiment. *14. Internationales Stuttgarter Symposium: Automobil-und Motorentechnik*.

Titsias, M. (2009). Variational learning of inducing variables in sparse Gaussian processes. *AISTATS*.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer New York.

Wang, K., Pleiss, G., Gardner, J., Tyree, S., Weinberger, K. Q., and Wilson, A. G. (2019). Exact Gaussian processes on a million data points. *NeurIPS*.

Wilson, A. and Nickisch, H. (2015). Kernel interpolation for scalable structured Gaussian processes (KISS-GP). *ICML*.

Zimmer, C., Driess, D., Meister, M., and Duy, N.-T. (2020). Adaptive discretization for evaluation of probabilistic cost functions. *AISTATS*.

Zimmer, C., Meister, M., and Nguyen-Tuong, D. (2018). Safe active learning for time-series modeling with Gaussian processes. *NeurIPS*.

Checklist

1. For all models and algorithms presented, check if you include:

(a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. **Yes**, see section 4.

(b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. **Yes**, see section 4.

(c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. **Yes**, see supplementary.

2. For any theoretical claim, check if you include:

   (a) Statements of the full set of assumptions of all theoretical results. **Yes**

   (b) Complete proofs of all theoretical results. **Yes**, see the cited references in section 4 and supplementary.

   (c) Clear explanations of any assumptions. **Yes**, see section 4.

3. For all figures and tables that present empirical results, check if you include:

   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). **Yes**, see supplementary.

   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). **Yes**, see supplementary.

   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). **Yes**, see captions of the respective figures.

   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). **Yes**, see supplementary.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:

   (a) Citations of the creator If your work uses existing assets. **Yes**, see supplementary.

   (b) The license information of the assets, if applicable. **Yes**, see supplementary.

   (c) New assets either in the supplemental material or as a URL, if applicable. **Yes**, see supplementary.

   (d) Information about consent from data providers/curators. **Not Applicable**, since we use synthethic data of publicly available models.

   (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. **Not Applicable**, since we use synthethic data of publicly available models.

5. If you used crowdsourcing or conducted research with human subjects, check if you include:

   (a) The full text of instructions given to participants and screenshots. **Not Applicable**

   (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. **Not Applicable**

   (c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. **Not Applicable**

Efficiently Computable Safety Bounds for Gaussian Processes in Active Learning
Supplementary Materials

1 Discrete and continuous trajectories

The theory and methodology developed in this paper is formulated in terms of continuous-time paths of Gaussian processes Z_t, parameterized by some trajectory $\tau(t) \in \mathbb{R}^n$, for $t \in T \subset [0,1]$. The set T is introduced to account for a discretization of the trajectory. Here, we highlight that we could instead describe the discretization by keeping $T = [0,1]$ and considering the alternative trajectory

$$\tau(t) = \begin{cases} \tau_1, & t \in [0, t_2), \\ \tau_i, & t \in [t_i, t_{i+1}), \quad i = 1, \dots, m-1 \\ \tau_m, & t \in [t_m, 1]. \end{cases}$$

Here, $\tau_i \in \mathbb{R}^n$ are the discrete points of the trajectory, and $0 \leq t_1 \leq \tau_2 \leq \ldots \leq \tau_m \leq 1$ are breakpoints of the step function. As Z_t is the posterior GP of $f(\tau(t))$, its paths are also step functions, and it holds that

$$\inf_{t \in [0,1]} Z_t = \min_{j=1,\dots,m} Z_{t_j}, \qquad \text{and} \qquad P^*(\tau) = P\left(\min_{j=1,\dots,m} Z_{t_j} \leq 0 \right).$$

Thus, all of our results readily transfer to finite discretizations of the trajectories τ. In particular, our methods are applicable for the discretizations chosen in Section 5.

2 Proofs

2.1 Proof of Theorem 1

Observe that $M \cdot \widehat{P}_{\mathrm{MC}}(M, \tau)$ admits a Binomial distribution with M trials and success probability $p = P^*(\tau)$. Moreover, let Y_r be a binomially distributed random variable with M_r trials and success probability α, and set $X = Y_r / M_r$. First, consider the case $p \geq \alpha$, such that in particular X_r is stochastically smaller than $\widehat{P}_{\mathrm{MC}}(M_r, \tau)$. We now use Okamoto's exponential bounds (Okamoto, 1959, Thm. 2) for the Binomial distribution: As $\alpha \leq 1/2$, for each $z > 0$ it holds $Q(X_r - \alpha \leq -z) \leq \exp(-\frac{M_r z^2}{2\alpha(1-\alpha)})$. Thus, the union bound yields

$$\begin{aligned} Q\left(\exists r \in \mathbb{N} : \widehat{P}^{+}_{\mathrm{MC}}(\tau, M_r, r, \epsilon, \alpha) < \alpha\right) &\leq \sum_{r=1}^{\infty} Q\left(\widehat{P}_{\mathrm{MC}}(M_r, \tau) < \alpha - \sqrt{\alpha(1-\alpha)} c_r\right) \\ &\leq \sum_{r=1}^{\infty} Q\left(X_r - \alpha < -\sqrt{\alpha(1-\alpha)} c_r\right) \\ &\leq \sum_{r=1}^{\infty} \exp\left(-\frac{M_r c_r^2}{2}\right) \\ &= \sum_{r=1}^{\infty} \frac{6\epsilon}{\pi^2 r^2} = \epsilon. \end{aligned}$$

For the case $p \le \alpha \le \frac{1}{2}$, the random variable X_r is stochastically larger than $\widehat{P}_{\mathrm{MC}}(M_r, \tau)$. Hence, we obtain

$$\begin{aligned}
Q\left(\exists r \in \mathbb{N} : \widehat{P}^{-}_{\mathrm{MC}}(\tau, M_r, r, \epsilon, \alpha) > \alpha\right) &\le \sum_{r=1}^{\infty} Q\left(\widehat{P}_{\mathrm{MC}}(M_r, \tau) > \alpha + \frac{c_r^2}{4} + c_r\sqrt{\alpha}\right) \\
&= \sum_{r=1}^{\infty} Q\left(\widehat{P}_{\mathrm{MC}}(M_r, \tau) > \left(\sqrt{\alpha} + \frac{c_r}{2}\right)^2\right) \\
&\le \sum_{r=1}^{\infty} Q\left(X_r > \left(\sqrt{\alpha} + \frac{c_r}{2}\right)^2\right).
\end{aligned}$$

Now we use another bound of Okamoto (Okamoto, 1959, Thm. 3): $Q(X_r > (\sqrt{\alpha} + z)^2) \le \exp(-2M_r z^2)$. Hence,

$$\begin{aligned}
Q\left(\exists r \in \mathbb{N} : \widehat{P}^{-}_{\mathrm{MC}}(\tau, M_r, r, \epsilon, \alpha) > \alpha\right) &\le \sum_{r=1}^{\infty} Q\left(X_r > \left(\sqrt{\alpha} + \frac{c_r}{2}\right)^2\right) \\
&\le \sum_{r=1}^{\infty} \exp\left(-\frac{M_r c_r^2}{2}\right) \\
&= \sum_{r=1}^{\infty} \frac{6\epsilon}{\pi^2 r^2} \qquad = \epsilon.
\end{aligned}$$

This completes the proof.

2.2 Proof of Theorem 5

Suppose that $P^{\dagger}(\tau) \ge \alpha$. Our choice of β_+ ensures that

$$\begin{aligned}
& Q\left(\tilde{m} \le q_{\beta_+, M_r}\right) \ge \chi(r, \epsilon) = 1 - \frac{6\epsilon}{\pi^2 r^2} \\
\iff\ & Q\left(\tilde{m} > q_{\beta_+, M_r}\right) \le \frac{6\epsilon}{\pi^2 r^2}.
\end{aligned}$$

Hence,

$$\begin{aligned}
Q\left(\widehat{P}^{\dagger}_{+}(M_r, r, \epsilon) < \alpha\right) &= Q\left(1 - \Phi\left(\tfrac{1 - q_{\beta_+, M_r}}{\tilde{\sigma}}\right) < \alpha \le P^{\dagger}(\tau)\right) \\
&\le Q\left(1 - \Phi\left(\tfrac{1 - q_{\beta_+, M_r}}{\tilde{\sigma}}\right) < P^{\dagger}(\tau)\right) \\
&= Q\left(1 - \Phi\left(\tfrac{1 - q_{\beta_+, M_r}}{\tilde{\sigma}}\right) < 1 - \Phi\left(\tfrac{1 - \tilde{m}}{\tilde{\sigma}}\right)\right) \\
&= Q\left(q_{\beta_+, M_r} < \tilde{m}\right) \\
&\le \frac{6\epsilon}{\pi^2 r^2}.
\end{aligned}$$

The union bound yields

$$Q\left(\exists r \in \mathbb{N} :\ \widehat{P}^{\dagger}_{+}(M_r, r, \epsilon) < \alpha\right) \ \le\ \sum_{r=1}^{\infty} Q\left(\widehat{P}^{\dagger}_{+}(M_r, r, \epsilon) < \alpha\right) \ \le\ \sum_{r=1}^{\infty} \frac{6\epsilon}{\pi^2 r^2} \ \le\ \epsilon.$$

This proves the first claim, and the second claim may be derived analogously.

2.3 Proof of Corollary 6

Suppose that $P^{*}(\tau) \ge \alpha$. By the definition of $\widehat{P}^{\diamond}_{+}(r, \epsilon, \alpha)$ and the union bound, we find that

$$\begin{aligned}
Q\left(\exists r \in \mathbb{N} :\ \widehat{P}^{\diamond}_{+}(r, \epsilon, \alpha) < \alpha\right) &\le Q\left(\exists r \in \mathbb{N} :\ \widehat{P}^{\dagger}_{+}(M_r, r, \tfrac{\epsilon}{2}) < \alpha\right) \\
&\quad + Q\left(\exists r \in \mathbb{N} :\ \widehat{P}^{+}_{\mathrm{MC}}(\tau, M_r, r, \tfrac{\epsilon}{2}, \alpha) < \alpha\right) \\
&\le \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2}.
\end{aligned}$$

The last inequality is due to Theorem 1 and Theorem 5, and establishes the first claim of Corollary 6. The second claim is identical to Theorem 1.

2.4 Further Proofs

The proof of Theorem 2 can be found in A.2.1 in van der Vaart and Wellner (1996). The proof of Theorem 4 is given in the main text as a combination of Theorem 2 and Remark 3.

3 Code

The code is provided under `github.com/joerntebbe/SafetyBounds4GPinAL` with the BSD-2 license. We use adapted code from Zimmer et al. (2018), which is provided under the MIT license, with a modified version of the Gaussian Process library from Rasmussen and Nickisch (2010) for MATLAB, which is provided under the FreeBSD license.
The experiments were carried out using CPU only with an AMD Ryzen 9 5950X driven at 3.4GhZ and 64GB RAM on MATLAB R2023a.

4 Algorithms

4.1 General implementation details

We use the active learning algorithm proposed in Zimmer et al. (2018) as baseline. Integrating our method into this algorithm comes with several challenges.
We have to add an immediate rejection of a candidate trajectory, if the mean has a changing sign, since our method is not applicable in this case. In order to provide meaningful gradients to the optimizer in this case, we implemented a heuristic to return a safety value which characterizes the trajectory as unsafe but also gives a metric on how far away the trajectory is from being safe without evaluating a particular bound. This results in a penalty term which is dependent on the distance from the mean to be classifiable by our method

$$P_{\text{unsafe}}(\tau_t) = 0.5 + \| \max(\mu_t, \mathbf{0}_m) \|_2$$

with $\mathbf{0}_m \in \mathbb{R}^m$ being a vector of zeros.
Moreover, if our method fails to make a decision with the provided budget of samples per safety evaluation, we declare the trajectory as unsafe. With the same purpose as before, we want to provide a numerical value which yields how far away from safe this trajectory is. In order to do so, we provide the lower bound of the confidence interval as the returning value.

As the sample sizes $M_1, \dots, M_R$ we use $M_1 = 100$ and double the size in each iteration, resulting in

$$M_r = 100 \cdot 2^{r-1}.$$

For the Himmelblau example we use $R = 14$, for the engine control example we use $R = 17$. These values were chosen based on prior experiments which observed the distribution of required samples to make a decision. In order to provide a fair comparison, i.e. the MC method respects the safety requirements, we chose the fixed sample size for the MC method to be M_{R-1}.

4.2 Adaptive Monte Carlo

Algorithm 1 provides Pseudocode for the adaptive Monte Carlo sampling scheme.

Algorithm 1 Adaptive Monte Carlo sampling

```
Require: Safety threshold: $\alpha > 0$, Threshold for confidence intervals $\epsilon > 0$,
  Discretization $t_1, \ldots, t_m$, sequence of sample sizes $0 = M_0 < M_1 < \ldots < M_R$, Posterior GP: $X_t$
  for $r = 1, \ldots, R$ do
      for $i = M_{r-1} + 1, \ldots, M_r$ do
          Simulate $(X_{t_j,i})_{j=1,\ldots,m}$
          $S_i \leftarrow \max_{j=1,\ldots,m} X_{t_j,i}$
      end for
      $\widehat{P} \leftarrow \frac{M_{r-1}}{M_r}\widehat{P} + \frac{1}{M_r}\sum_{i=M_{r-1}+1}^{M_r} \mathbf{1}(S_i > 1)$
      $\widehat{P}^{\pm}_{\mathrm{MC}} \leftarrow \widehat{P} \pm \sqrt{\frac{2\alpha(1-\alpha)}{M_r}\left|\log\frac{6\epsilon}{\pi^2 r^2}\right|}$
      if $\widehat{P}^{+}_{\mathrm{MC}} < \alpha$ then
          return SAFE
      else if $\widehat{P}^{-}_{\mathrm{MC}} > \alpha$ then
          return UNSAFE
      end if
  end for
  return UNSAFE
```

4.3 Adaptive Borell-TIS

Algorithm 2 provides Pseudocode for the safety evaluation using the Borell-TIS inequality and the adaptive sampling scheme of the median.

Algorithm 2 Adaptive sampling for Borell-TIS

```
Require: Safety threshold: $\alpha > 0$, Threshold for confidence intervals $\epsilon > 0$,
  Discretization $t_1, \ldots, t_m$, sequence of sample sizes $0 = M_0 < M_1 < \ldots < M_R$, Posterior GP: $X_t$
  for $r = 1, \ldots, R$ do
      $\beta_{\pm} \leftarrow \frac{1}{2} \pm \Phi^{-1}(1 - \chi(M_r, \epsilon))/\sqrt{4M_r}$
      for $i = M_{r-1} + 1, \ldots, M_r$ do
          Simulate $(X_{t_j,i})_{j=1,\ldots,m}$
          $S_i \leftarrow \max_{j=1,\ldots,m} X_{t_j,i}$
      end for
      $q_{\pm} \leftarrow q_{\beta_{\pm}(M_r,k,\epsilon),M_r}(S_1, \ldots S_{M_r})$
      $\widehat{P}^{\dagger}_{\pm} \leftarrow 1 - \Phi\left(\frac{1-q_{\pm}}{\sigma_m}\right)$
      if $\widehat{P}^{\dagger}_{+} \leq \alpha$ then
          return SAFE
      else if $\widehat{P}^{\dagger}_{-} \geq \alpha$ then
          return UNSAFE
      end if
  end for
  return UNSAFE
```

4.4 Hybrid scheme

Algorithm 3 provides Pseudocode for the safety evaluation using the adaptive hybrid scheme described in Section 4.4 of the article.

Algorithm 3 Adaptive hybrid scheme

```
Require: Safety threshold: $\alpha > 0$, Threshold for confidence intervals $\epsilon > 0$,
  Discretization $t_1, \ldots, t_m$, sequence of sample sizes $0 = M_0 < M_1 < \ldots < M_R$, Posterior GP: $X_t$
  for $r = 1, \ldots, R$ do
    $\beta_\pm \leftarrow \frac{1}{2} \pm \Phi^{-1}(1 - \chi(Mk, \epsilon))/\sqrt{4M_r}$
    for $i = M_{r-1} + 1, \ldots, M_r$ do
      Simulate $(X_{t_j,i})_{j=1,\ldots,m}$
      $S_i \leftarrow \max_{j=1,\ldots,m} X_{t_j,i}$
    end for
    $\widehat{P} \leftarrow \frac{M_{r-1}}{M_r}\widehat{P} + \frac{1}{M_r}\sum_{i=M_{r-1}+1}^{M_r} \mathbf{1}(S_i > 1)$
    $\widehat{P}^+_{\mathrm{MC}} \leftarrow \widehat{P} + \sqrt{\frac{2\alpha(1-\alpha)}{M_r}\left|\log \frac{3\epsilon}{\pi^2 r^2}\right|}$
    $\widehat{P}^-_{\mathrm{MC}} \leftarrow \widehat{P} - \sqrt{\frac{2\alpha(1-\alpha)}{M}\left|\log \frac{3\epsilon}{\pi^2 r^2}\right|}$          ▷ Critical values based on MC scheme
    $q_+ \leftarrow q_{\beta_+(M_r,k,\frac{\epsilon}{2}),M_r}(S_1, \ldots S_{M_r})$
    $\widehat{P}^\dagger_+ \leftarrow 1 - \Phi\left(\frac{1-q_\pm}{\tilde{\sigma_m}}\right)$          ▷ Critical value based on Borell-TIS scheme
    if $\min[\widehat{P}^\dagger_+, \widehat{P}^+_{\mathrm{MC}}] \leq \alpha$ then
      return SAFE
    else if $\widehat{P}^-_{\mathrm{MC}} \geq \alpha$ then
      return UNSAFE
    end if
  end for
  return UNSAFE
```

5 Further information on the examples

In this section we provide further information on the experiments presented in the main text. We observe the quantities $RMSE$ and c_h, as well as the quantities n_{SAL} which is the number of iterations of the active learning algorithm, and n_f, which is defined as the number of training points, that are unsafe due to the ground truth. A detailed discussion on this can be found in the respective subsections.

5.1 Himmelblau's function exploration

Table 1: Quantities for Himmelblau's function exploration:

method / α	n_{SAL}	RMSE	c_h	n_f
MC / 0.1	10.9 ± 1.7	0.2873 ± 0.0669	0.4906 ± 0.0540	0.0000 ± 0.0000
AMC (Ours) / 0.1	38.2 ± 7.4	0.1553 ± 0.0543	0.7724 ± 0.0630	0.2000 ± 0.6325
AB (Ours) / 0.1	**65.1 ± 7.4**	**0.1080 ± 0.0302**	**0.8621 ± 0.0326**	0.7000 ± 1.4944
ABM (Ours) / 0.1	59.6 ± 9.5	0.1511 ± 0.0801	0.8067 ± 0.0625	1.1000 ± 3.1429
MC / 0.01	10.5 ± 1.9	0.3356 ± 0.0906	0.4476 ± 0.0711	0.0000 ± 0.0000
AMC (Ours) / 0.01	26.3 ± 3.9	0.2194 ± 0.0798	0.6078 ± 0.0521	0.0000 ± 0.0000
AB (Ours) / 0.01	50.8 ± 5.1	**0.1745 ± 0.0812**	**0.7338 ± 0.0637**	0.0000 ± 0.0000
ABM (Ours) / 0.01	**70.1 ± 3.0**	0.1771 ± 0.0498	0.7332 ± 0.0569	0.0000 ± 0.0000
MC / 0.001	10.6 ± 2.3	0.3482 ± 0.0648	0.4352 ± 0.0634	0.0000 ± 0.0000
AMC (Ours) / 0.001	31.4 ± 6.0	0.2434 ± 0.0793	0.6457 ± 0.0556	0.0000 ± 0.0000
AB (Ours) / 0.001	59.1 ± 10.5	**0.1619 ± 0.0924**	**0.7895 ± 0.0631**	0.1000 ± 0.3162
ABM (Ours) / 0.001	**64.6 ± 6.4**	0.1738 ± 0.0934	0.7470 ± 0.0887	0.3000 ± 0.9487

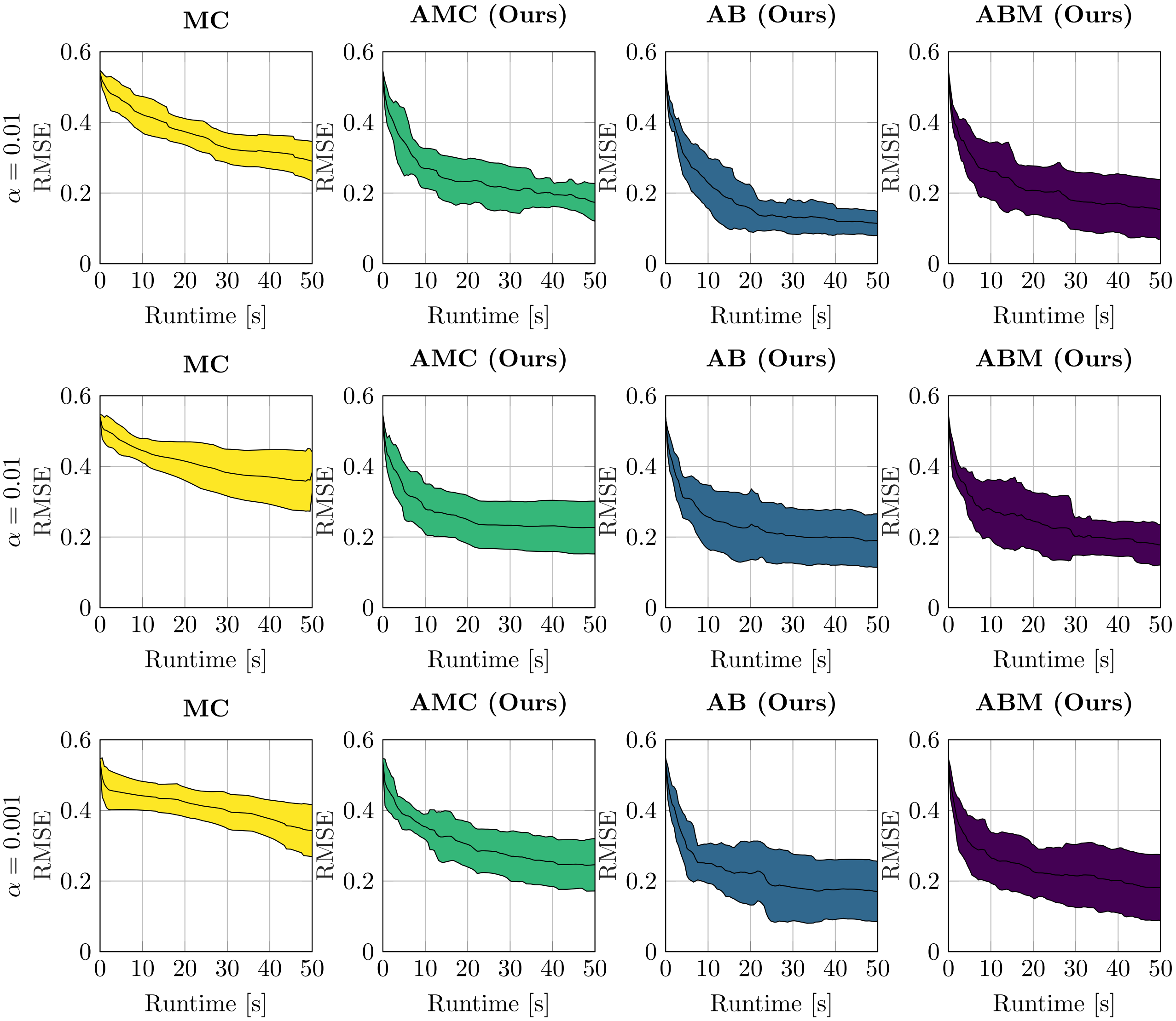


Figure 1: Himmelblau's function exploration: These plots show the RMSE for three different values of α and the four algorithms. The results are averaged over ten independent seeds and regions are 2σ confidence intervals.

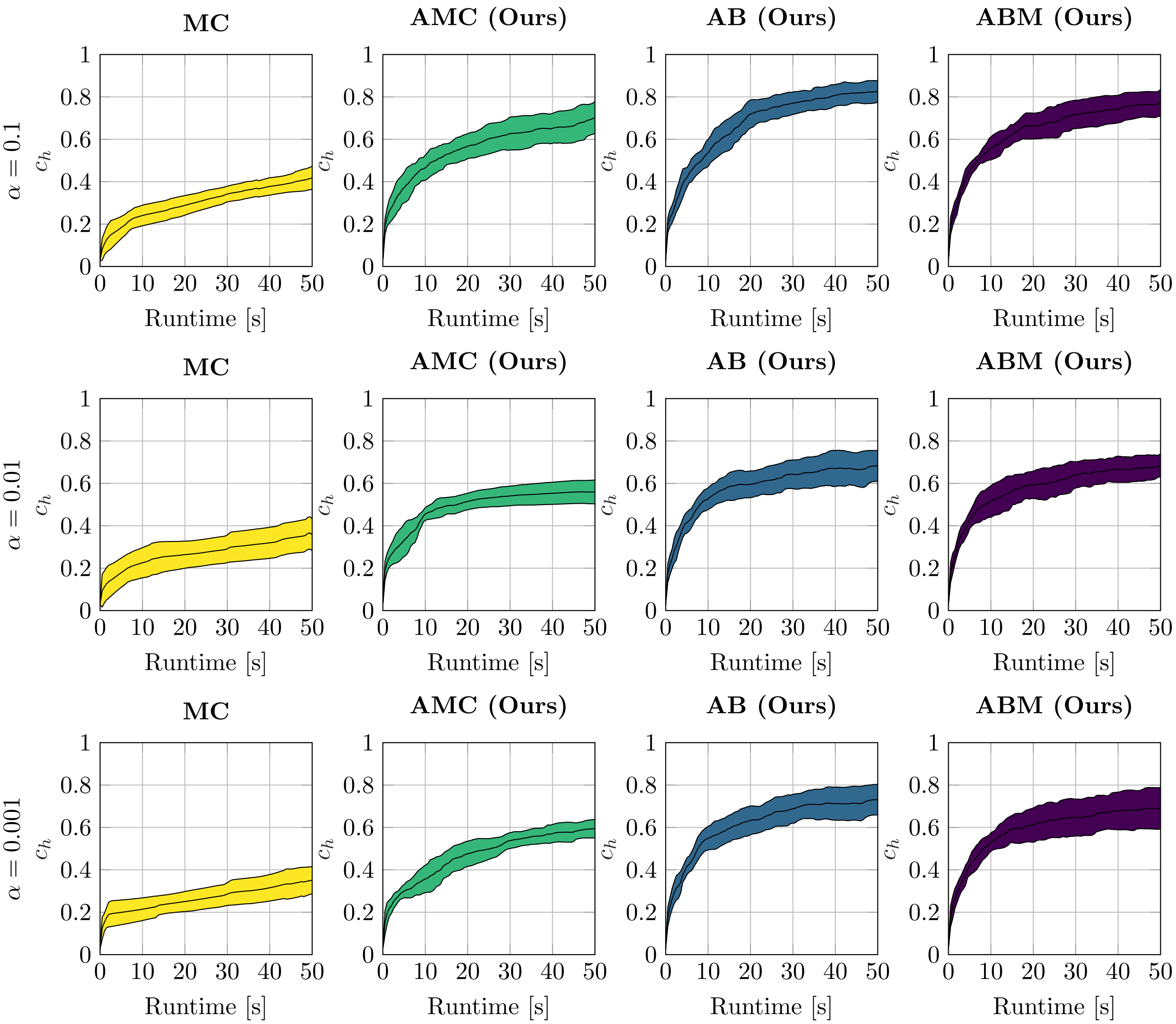


Figure 2: Himmelblau function exploration health coverage: These plots show the health coverage c_h for three different values of α and the four algorithms. The results are averaged over ten independent seeds and regions are 2σ confidence intervals.

We use fixed hyperparameters without optimization. The hyperparameters are $\ell_1^2 = \ell_2^2 = 1.0$, $\sigma_f^2 = 1$ and $\sigma_n = 0.01$. We scale the function with a factor of 0.01 and add normal distributed noise with zero mean and a standard deviation of 0.01 which coincides with σ_n. We discretize the trajectories with $m = 5$, but only use the endpoint of an explored trajectory as new measurement which is added to the training points. We present the additional results in Figures 1 and 2, and Table 1. We see, that our proposed algorithms using the (B.1) bound perform the best with AMC (Ours) also outperforming MC. This results in much more iterations made by the safe active learning algorithm due to fewer samples needed for safety evaluation.

5.2 Engine control

Table 2: Quantities for engine control

method / α	n_{SAL}	RMSE	c_h	n_f
MC / 0.1	12.7 ± 1.4	11.6250 ± 0.2191	0.5075 ± 0.0280	1.5000 ± 2.1213
AMC (Ours) / 0.1	74.8 ± 9.8	10.8945 ± 0.2759	0.6045 ± 0.0220	10.1000 ± 10.3220
AB (Ours) / 0.1	**111.0 ± 7.7**	**10.5050 ± 0.4379**	**0.6643 ± 0.0306**	16.7000 ± 11.7004
ABM (Ours) / 0.1	72.7 ± 9.2	10.8287 ± 0.3261	0.6262 ± 0.0300	9.1000 ± 8.5823
MC / 0.01	12.2 ± 1.7	11.6765 ± 0.2750	0.5007 ± 0.0358	1.1000 ± 1.4491
AMC (Ours) / 0.01	98.1 ± 7.2	11.2986 ± 0.1837	0.5815 ± 0.0253	6.6000 ± 4.4771
AB (Ours) / 0.01	129.3 ± 4.4	10.9095 ± 0.2787	0.6258 ± 0.0215	5.3000 ± 2.8304
ABM (Ours) / 0.01	**135.9 ± 3.0**	**10.8500 ± 0.2597**	**0.6387 ± 0.0289**	15.0000 ± 13.6870
MC / 0.001	12.5 ± 1.7	11.7215 ± 0.2253	0.4982 ± 0.0333	1.2000 ± 1.5492
AMC (Ours) / 0.001	94.9 ± 7.1	11.5351 ± 0.2537	0.5600 ± 0.0356	6.1000 ± 6.7897
AB (Ours) / 0.001	131.1 ± 2.1	11.0656 ± 0.3132	0.6111 ± 0.0361	7.8000 ± 4.0222
ABM (Ours) / 0.001	**134.3 ± 2.8**	**10.9203 ± 0.3193**	**0.6161 ± 0.0240**	10.8000 ± 6.7626

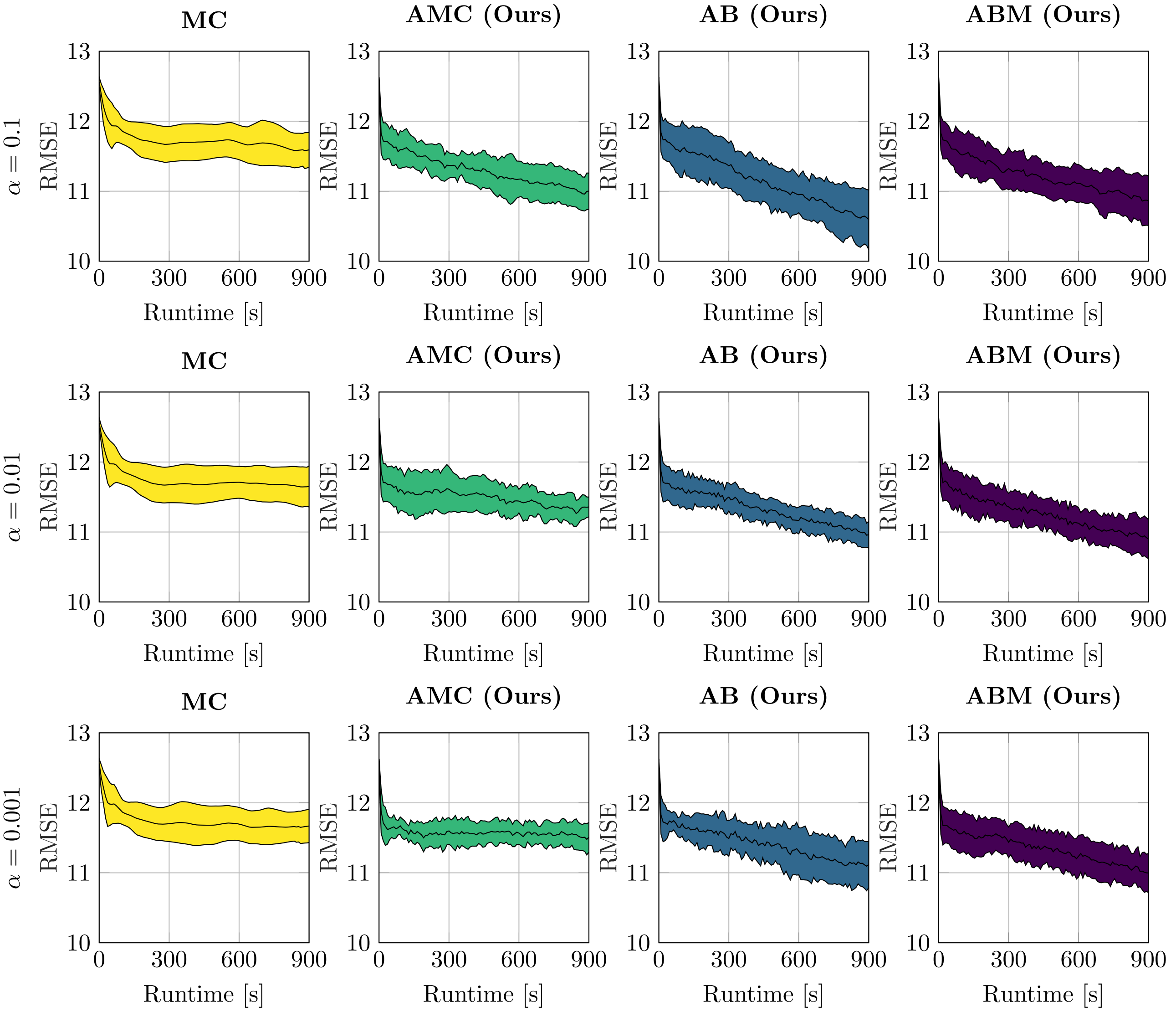


Figure 3: Engine control RMSE: These plots show the RMSE for three different values of α and the four algorithms. The results are averaged over ten independent seeds and regions are 2σ confidence intervals.

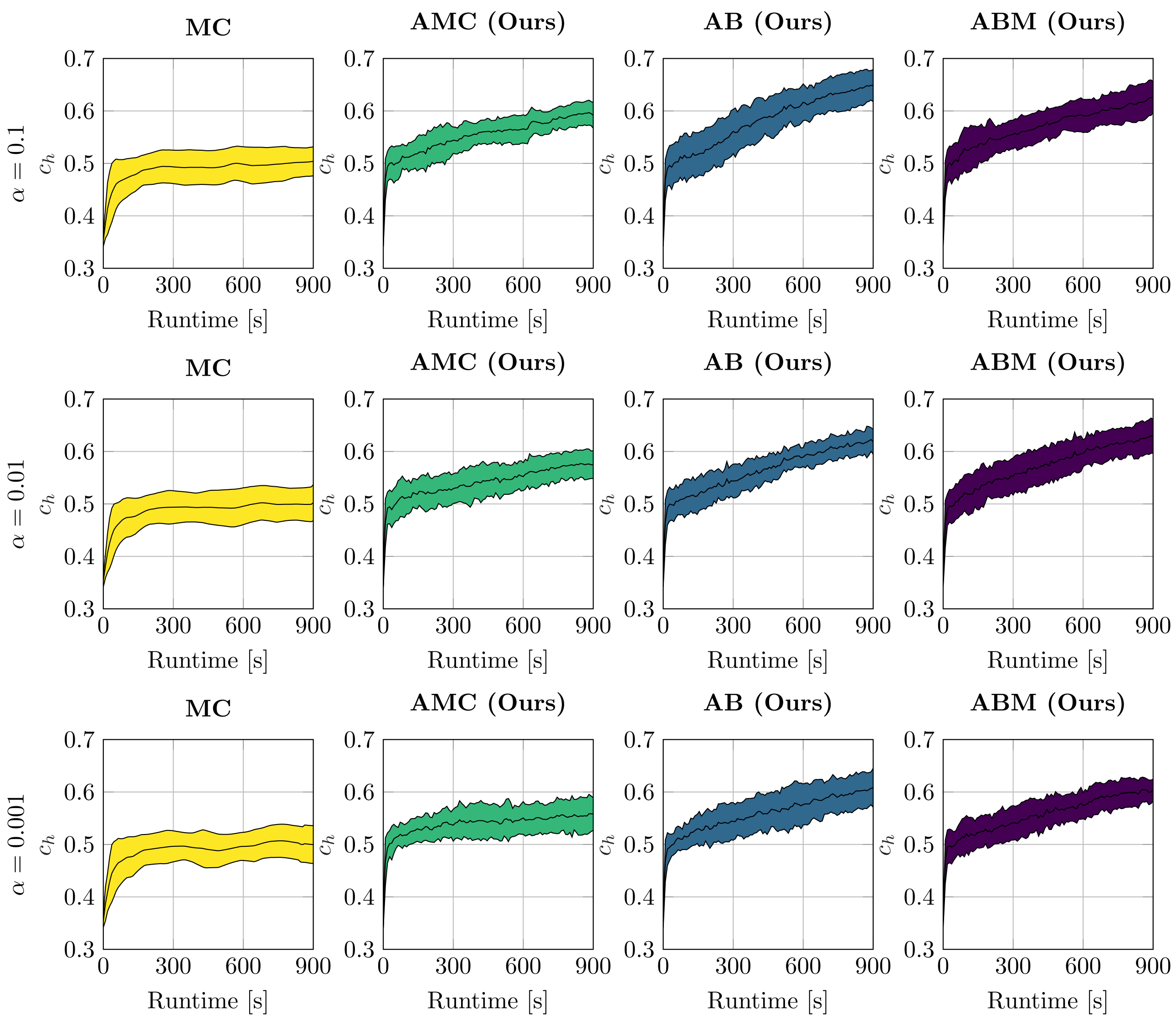


Figure 4: Engine control health coverage: These plots show the health coverage c_h for three different values of α and the four algorithms. The results are averaged over ten independent seeds and regions are 2σ confidence intervals.

We perform constrained hyperparameter optimization in each iteration. The initial hyperparameters, as well as the lower and upper bounds for the constrained optimization are taken from Zimmer et al. (2018). We use a discretization of $m = 5$ and add each of these points to the training points after exploration. Furthermore we add another heuristic to the algorithms containing the B.1 bound. If we cannot compute the lower bound for the confidence interval, since we are out of the feasible interval, we classify the point as unsafe.

We present further results on the experiments of the main text in Figure 3 and 4, and Table 2. The number of unsafe training points n_f is significantly higher than for the first example. We explain this with the significant higher modeling error of the GP which can also be seen in the RMSE. Since our safety estimation relies on the GP, these errors cannot be avoided in the generation of safety bounds for dynamic systems.

References

Okamoto, M. (1959). Some inequalities relating to the partial sum of binomial probabilities. *Annals of the Institute of Statistical Mathematics*, 10(1):29–35.

Rasmussen, C. E. and Nickisch, H. (2010). Gaussian processes for machine learning (gpml) toolbox. *The Journal of Machine Learning Research*, 11:3011–3015.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer New York.

Zimmer, C., Meister, M., and Nguyen-Tuong, D. (2018). Safe active learning for time-series modeling with Gaussian processes. *NeurIPS*.