Learning how to safely explore dynamic systems

Christoph Zimmer CHRISTOPH.ZIMMER@DE.BOSCH.COM
Bosch Center for Artificial Intelligence, Robert Bosch GmbH
Robert-Bosch-Campus 1
71272 Renningen, Germany

Cen-You Li CEN-YOU.LI@DE.BOSCH.COM
Bosch Center for Artificial Intelligence, Robert Bosch GmbH
Robert-Bosch-Campus 1
71272 Renningen, Germany

Mark Schillinger MARK.SCHILLINGER@DE.BOSCH.COM
Bosch Engineering GmbH
Robert-Bosch-Allee 1
74232 Abstatt, Germany

Mona Meister MONA.MEISTER@DE.BOSCH.COM
Bosch Center for Artificial Intelligence, Robert Bosch GmbH
Robert-Bosch-Campus 1
71272 Renningen, Germany

Barbara Rakitsch BARBARA.RAKITSCH@DE.BOSCH.COM
Bosch Center for Artificial Intelligence, Robert Bosch GmbH
Robert-Bosch-Campus 1
71272 Renningen, Germany

Duy Nguyen-Tuong DUY.NGUYEN-TUONG@DE.BOSCH.COM
Bosch Center for Artificial Intelligence, Robert Bosch GmbH
Robert-Bosch-Campus 1
71272 Renningen, Germany

Abstract

Computational modeling is central in gaining systems' understanding in a broad range of domains. Data based modeling has become an essential part of these efforts. While in some domains labelled data is abundant, e.g., search engines queries, in other domains labeling is time consuming, e.g., medical image analysis, comes at high financial cost and sometimes is even safety critical, e.g., experiments in chemical reactors or test benches. Therefore, approaches for automatic data generation under safety constraints are a key enabler for automatic data generation in those systems. We present approaches for safe learning and combine it with dynamic systems, systems with multiple outputs as well as real world industrial applications.

Keywords: Safe learning, Active Learning, automatic labeling, experimental design, Gaussian processes

1. Introduction

Safe exploration is highly relevant in domains like robotics (Sui et al., 2018; Berkenkamp et al., 2016; Baumann et al., 2021), energy management (Galichet et al., 2013), terrain exploraion (Moldovan and Abbeel, 2012; Turchetta et al., 2019) or engine modeling (Schreiter et al., 2015; Zimmer et al., 2018). If a safe area of operation is known a-priori by e.g., domain experts, this can be easily incorporated as a constraint. The more common and challenging scenario are unkown safe areas. Here, we focus on this aspect of how data for machine learning models can be safely generated when the safety constraints have to be learned on the fly during the course of the measurements.

Commonly, it is assumed that during the experiment, some feedback on the system's safety or health can be obtained (Schreiter et al., 2015; Zimmer et al., 2018; Turchetta et al., 2019; Baumann et al., 2021). This feedback can be used to learn a safety model. Gaussian processes are a favorable choice as they are a flexible modeling tool and also provide an uncertainty quantification.

Safe learning has been combined with Bayesian optimization (Berkenkamp et al., 2016), Active Learning (Schreiter et al., 2015) or dynamic Active Learning (Zimmer et al., 2018). The meaning of safety or system's health varies from crashing a drone or blowing up an engine over some moderate harm as scratches to rather comfort-related issues as system shutdowns. Therefore, it is beneficial to relate the potential harm and the strictness at which the safety constraints are adhered to (Berkenkamp et al., 2016; Zimmer et al., 2018).

In this work, we contextualize recently published methods on safe Active Learning and their real world applications.

2. Gaussian Processes

Gaussian processes (GP) are a widely used data based and probabilistic modeling technique. Given some data set $D = \{x_i, y_i\}_{i=1}^n$ consisting of inputs x_i and labels y_i, a mean, and a kernel function k, the GP returns a predictive distribution for novel points x^* of interest with mean $\mu(x^*) = k_n^T(k_{nn}+\sigma^2 I)^{-1}Y$ and covariance function $\Sigma(x^*) = k(x^*, x^*) - k_n^T(k_{nn}+\sigma^2 I)^{-1}k_n$ where $k_n = k(X, x^*)$, $k_{nn} = k(X, X)$, $X = (x_1, \ldots, x_n)^T$, $Y = (y_1, \ldots, y_n)^T$ as well as I is an identiy matrix and σ^2 is a noise level, see (Rasmussen and Williams, 2006) and (Quiñonero-Candela and Rasmussen, 2005; Snelson and Ghahramani, 2006; Titsias, 2009) for sparse approximation in case of large data sets.

3. Safe Learning

We are interested in safely learning a data based model for a regression task. We assume that aside of the regression variable y there is a possibly different safety indicator z that can be measured during a measurement campaign. z could be for example a pressure in an engine. We furthermore assume that domain experts can identify which levels of the safety indicator are safe and denote these by S. While domain experts are usually able to provide this information in the (safety) output space Z, the safe part of the input space X

is usually unkown. In the example of the pressure, domain experts might know a critical maximal pressure z_{max} and, therefore, $S = \{z < z_{max}\}$.

From now on, we use two GP models: one for the regression task, denoted by μ_f and Σ_f, and one for the safety indicator, denoted by μ_s and Σ_s. Given a new candidate for measurement x^*, the safety GP yields us the probability that the candidate is safe:

$$\xi(x^*) = \int_{z \in S} N(z|\mu_s(x^*), \Sigma_s(x^*))dz \tag{1}$$

If the probability of x^* to be safe is higher than a user defined threshold $0 < \alpha \leq 1$, then measurements y^* and z^* for x^* are conducted. The threshold α depends on the severity of a potential violation. In case of crashes or explosion, one might choose an α close to one, while in cases of mainly unwanted shutdowns one could choose α further away from one.

If more than one point x^* is to be selected, the GP predictive distribution (equation 1) is multi-variate, rendering the problem analytically intractable due to the integration over a multivariate Gaussian. While simple Monte-Carlo is an approach, expectation propagation (Genz, 1992) or adaptive discretization (Zimmer et al., 2020) can be more sophisticated answers.

Above, the safety condition has been defined as a regression task. While a classification of safety probability is also possible (Schreiter et al., 2015), our approach is aware of the distance to the critical level and approaches the safety boundary slowly (Figure 1).

3.1 Safe Active Learning

Test bench time is scarce and valuable so experimenters want to get the most out of it. Active Learning (AL) is a form of sequential experimental design that chooses information optimal points given some preceding model, updates the model, and continues by choosing the next set of information optimal points.

Combining Safe with Active Learning means combining the information maximization with a safety constraint, mathematically leading to a constrained optimization problem for new points x^*:

$$\begin{aligned} x^* &= \text{argmax}_x \mathcal{I}(x) \\ &s.t. \quad \xi(x) > 1 - \alpha \end{aligned}$$

with an information criterion based on e.g., entropy or mutual information. While mutual information has benefits (Krause et al., 2008), in case of GP entropy is easy to compute as it is proportional to the predictive co-variance and, therefore, $\mathcal{I}(x) = det(\Sigma_f(x))$. As computational speed matters in real world applications, we focus on entropy in the following.

3.2 Dynamic Safe Active Learning

Some applications require dynamic modeling, e.g., the same speed of an engine might lead to different pressures depending on whether one is speeding up or slowing down. Nonlinear

autoregressive GPs model such dynamic behavior and parametrized trajectories containing $m+1$ discretization points $\tau(\eta) = (x_0, x_1, \ldots, x_m = \eta)$ can explore the space (Zimmer et al., 2018). η could, for example, contain start and end point of a ramp (see A1).

3.3 Safe Active Learning for multi-output

Many systems allow simultaneously measuring multiple outputs. Multi-output GPs (MOGP) exploit the correlation between the output to increase model accuracy. (Li et al., 2022) showed that Safe AL for MOGP (SAL-MOGP) exploits the correlation leading to an even more efficient data generation. The correlation also allows for non-synchronous data measurement of different outputs. The model can thus explore individual data domains for each output, which provides further flexibility in the measurement processes.

4. Applications

4.1 A toy example

We first illustrate how the safe area is learned on the fly. We start with a few meaurements in a small rectangular area that we assume can be determined as safe by domain experts. Then, our safe AL technology automatically decides for new safe measurments and upon receiving labels of the safety indicator extends the knowlegde of the safe area.

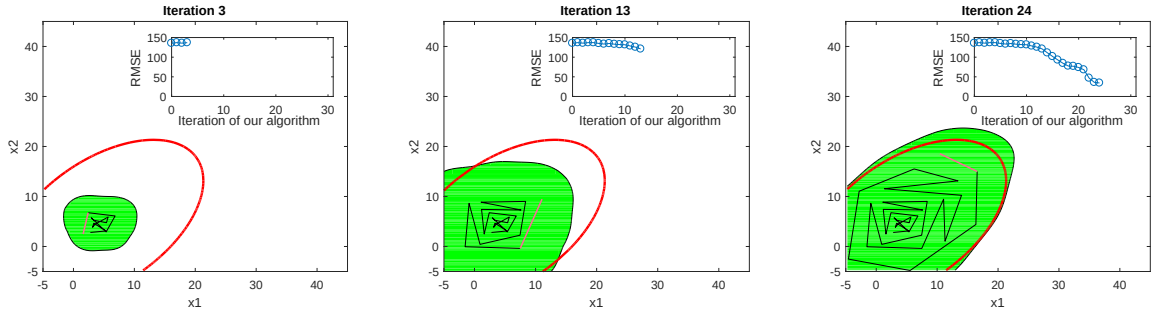

Figure 1: **Illustration how the safe area is learned on the fly.** Left: based on initial data (black line) a safety GP is trained. With its help the green area that is safe with high probability can be determined. Middle: a new measuerment is selected (purple) and based on its labels, the safety GP updated (green area becomes bigger). Right: with more measurements the safety GP is able to correctly identify the safe area. Inlay figures: test RMSE of regression problem is decreasing. Figure from (Zimmer et al., 2018).

4.2 High Pressure Fuel Supply System modeling

When calibrating combustion engines, models are used to speed up the calibration process, reduce the risk of damaging expensive prototypes, and reduce the time the real system needs to be available. One example is the high pressure fuel supply system of a gasoline engine. When considering no-load operation, this system features two inputs (engine speed and fuel pump actuation) and one output (rail pressure), which must not exceed a given limit for safe operation.

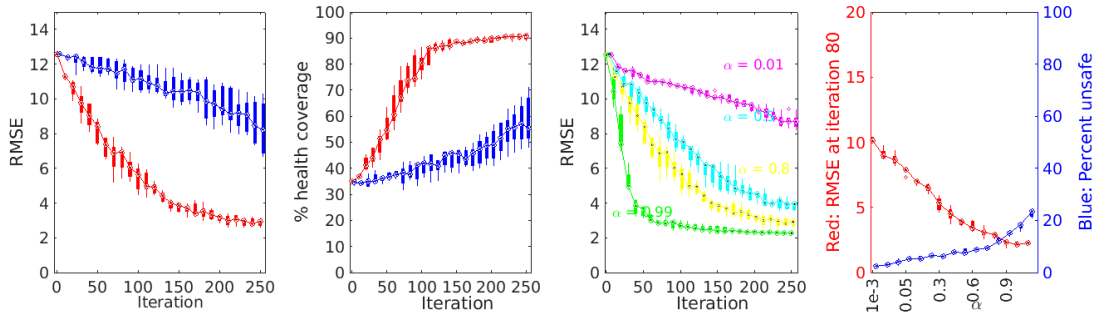

Figure 2: The first two pictures from the left show the comparison of the SAL-NX (red line) with random selection with safe constraints (blue line), with respect to model approximation (lower RMSE is better) and coverage of safe regions (higher coverage is better). The last two pictures show the impact of the safety threshold on the approximation error, and failures during exploration.

We first consider a simulator for this high pressure fuel supply system that allows us to carry out various analyses how quickly the safe area is learned (Figure 2 middle left), the effect of the safety value α (Figure 2 middle right and right) as well as AL (Figure 2 left). Next, we demonstrate that the dynamic safe AL algorithm (DSAL) can be successfully and safely used to conduct safe and information optimal experiments at the high pressure fuel supply system in a real car (Figure 3).

Figure 4 left shows the excitation signals created by DSAL. Note that the maximum pressure of 18 MPa is never exceeded. Figure 4 right compares the modeling performance of DSAL with space-filling ramp and chirp signals on different test data sets. DSAL always performs on par or better than the other signal type, which does not directly match the test signal type. The modeling error is comparatively high, because a very simple NX structure was used, which is not sufficient for this system. See (Schillinger, 2019) for more details.

Figure 3: **Dynamic safe Active Learning steering a car's engine to conduct measurements for a railpressure system.** Experimental set-up connecting the car's engine to the Matlabe based algorithm.

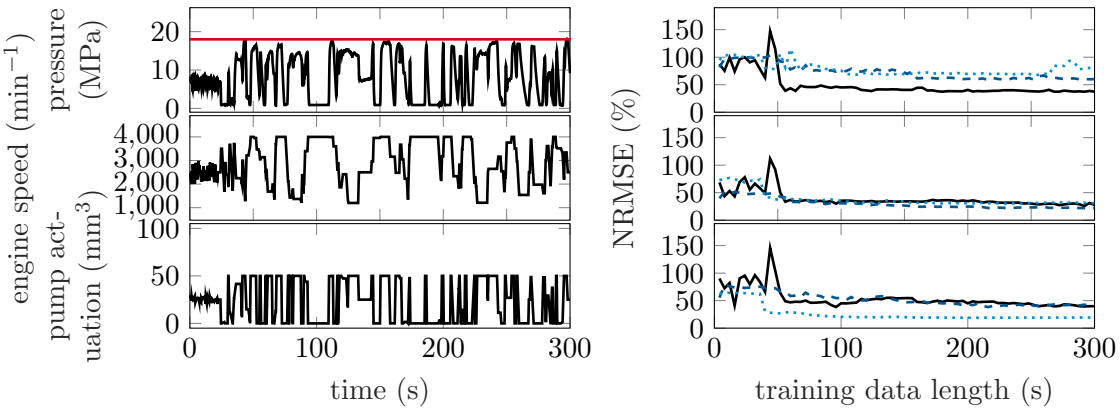

Figure 4: Left: Output (top) and inputs (middle and bottom) for a dynamic SAL measurements at a high pressure fuel supply system. The maximum allowed pressure was set to 18 MPa and never exceeded.
Right: normalized root mean square error on three test datasets with three different excitation signals. Top: DSAL data as test data, middle: ramps as test data, bottom: chirps as test data. Solid line: DSAL training data, dashed line: ramp training data, dotted line: chirp training data. Images from Schillinger (2019).

4.3 Emission modeling

The role of virtual sensors in the automotive industry has been growing tremendously over the past years and their applications can be found in numerous disciplines, such as engine calibration, diagnostic tools or validation tests. Here, we are concerned with building a virtual sensor over the emissions of a car's engine while keeping the total number of measurements low[1]. More concretely, we aim at modeling the outputs channels hydrocarbon (HC) and oxygen (O2) simultaneously while considering it safety critical that the temperature of the engine stays below a certain threshold. Since HC and O2 are known to be physically related, our SAL-MOGP framework is highly suitable for this task. We compare our approach with conventional safe AL or MOGP without AL in Figure A1. Our results confirm that combining AL with MO learning achieves a smaller test error when keeping the number of samples fixed.

5. Outlook

We believe in the relevance of real world Active Learning. Many challenges will come up and their solution will lead to further academic results and *the real world* impact. To support the progress of the field, relatistic simulators or even better collaborative access to test benches are needed to allow for scientifically accurate benchmarking. The interface between test bench and algorithm is a (more engineering related) challenge that needs standardized answers to scale the research communities methods to wide applicability.

1. Data is available at https://github.com/boschresearch/Bosch-Engine-Datasets/tree/master/gengine1

Acknowledgments

This work was supported by Bosch Center for Articial Intelligence, which provided financial support, computers and GPU clusters. The Bosch Group is carbon neutral. Administration, manufacturing, and research activities do no longer leave a carbon footprint. This also includes GPU clusters on which some of the experiments have been performed.

References

Dominik Baumann, Alonso Marco, Matteo Turchetta, and Sebastian Trimpe. Gosafe: Globally optimal safe robot learning. In *2021 IEEE International Conference on Robotics and Automation (ICRA 2021)*, pages 4452–4458, Piscataway, NJ, October 2021. IEEE. doi: 10.1109/ICRA48506.2021.9560738.

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

N. Galichet, M. Sebag, and O. Teytaud. Exploration vs Exploitation vs Safety: Risk-Aware Multi-Armed Bandits. In *Proceedings of the 5th Asian Conference on Machine Learning*, 2013.

Alan Genz. Numerical computation of multivariate normal probabilities. *Journal of Computational and Graphical Statistics*, 1:141–150, 1992.

Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research*, 2008.

Cen-You Li, Barbara Rakitsch, and Christoph Zimmer. Safe active learning for multi-output gaussian processes. In *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, 2022.

T. M. Moldovan and P. Abbeel. Safe Exploration in Markov Decision Processes. In *Proceedings of the 29th International Conference on Machine Learning*, 2012.

Joaquin Quiñonero-Candela and Carl Edward Rasmussen. A Unifying View of Sparse Approximate Gaussian Process Regression. In *Journal of Machine Learning Research*, 2005.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2006.

Mark Schillinger. *Safe and Dynamic Design of Experiments*. Phd thesis, University of Siegen, 2019.

Jens Schreiter, Duy Nguyen-Tuong, Mona Eberts, Bastian Bischoff, Heiner Markert, and Marc Toussaint. Safe Exploration for Active Learning with Gaussian Processes. In *ECML/PKDD*, volume 9286, 2015.

Edward Lloyd Snelson and Zoubin Ghahramani. Sparse Gaussian Processes using Pseudo-inputs. In *Advances in Neural Information Processing Systems*, 2006.

Yanan Sui, Vincent Zhuang, Joel W. Burdick, and Yisong Yue. Stagewise safe bayesian optimization with gaussian processes. In *35th International Conference on Machine Learning*, 2018.

Michalis K. Titsias. Variational Learning of Inducing Variables in Sparse Gaussian Processes. In *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*, 2009.

Matteo Turchetta, Felix Berkenkamp, and Andreas Krause. Safe exploration for interactive machine learning. In H. Wallach, H. Larochelle, A. Beygelzimer, F. dAlché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL `https://proceedings.neurips.cc/paper/2019/file/4f398cb9d6bc79ae567298335b51ba8a-Paper.pdf`.

Christoph Zimmer, Mona Meister, and Duy Nguyen-Tuong. Safe active learning for time-series modeling with gaussian processes. In *32nd Conference on Neural Information Processing Systems*, 2018.

Christoph Zimmer, Danny Dries, Mona Meister, and Duy Nguyen-Tuong. Adaptive Discretization for Probabilistic Safety Cost Functions Evaluation. In *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*, 2020.

Appendix

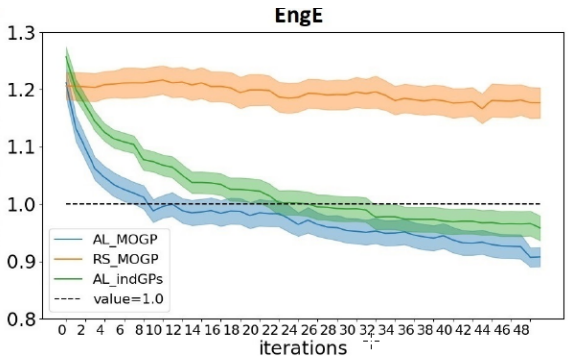

Figure A1: **SAL for MOGP even more data efficient.** The test error of SAL for MOGP considering correlation between outputs decreases even faster than those for conventional SAL or MOGP without AL.

Appendix A1. Selection of new points in the dynamic scenario

As mentioned in section 3.2 we use GPs with a nonlinear autoregressive structure. Denoting the controls of the test bench - the variables that we can choose - with u, each input x of the GP consists of the current and past values of u as well as possibly in case of output feedback also past values of the output y.

Let q denote the length of the considered output history and d the length of the considered history of u. Then, an input point takes the following form:
$x_k = (y_{k-1}, \ldots, y_{k-q}, u_k, u_{k-1}, \ldots, u_{k-d})$.

In this work, we consider the case of $q = 0$ and $d > 0$ (no output feedback). We plan trajectories of $m+1$ points $\tau(\eta) = (x_0, x_1, \ldots, x_m = \eta)$ with x_k as above and appropriately choose the parameter η. We denote $\tau_x = (x_1, \ldots, x_m = \eta)$ as a trajectory without x_0.

In the case of ramps, the parameter η could simply be the end point of the ramp (u_k). Now, we determine η similar to section 3.1 by solving a constrained optimization problem:

$$\begin{aligned} \eta^* &= \operatorname{argmax}_\eta \mathcal{I}(\tau_x(\eta)) \\ &s.t. \quad \xi(\tau(\eta)) > 1 - \alpha \end{aligned}$$

Note that we use τ_x for the information gain as the information from x_0 has already been counted by the previous trajectory. We use τ for the safety as the piece between x_0 (last point of previous trajectory) and x_1 should also be checked for safety.

Note that in the case of $q > 0$ (output feedback), calculating the information (predictive variance) of the m^{th} point of a trajetory requires the output value y_{m-1} which has not yet been observed as it is in the future. One can use the predictive mean of the GP as a surrogate for this output value y_{m-1} and others in the future. While this would ignore the uncertainty of the predicted y_{m-1}, moment matching methods or sampling could be used for a proper consideration of this uncertainty. Note that the input excitation is still performed through the manipulated variable u_k even in the case of output feedback.