Safe Controller Optimization for Quadrotors with Gaussian Processes

Felix Berkenkamp, Angela P. Schoellig, and Andreas Krause

***Abstract*— One of the most fundamental problems when designing controllers for dynamic systems is the tuning of the controller parameters. Typically, a model of the system is used to design an initial controller, but ultimately, the controller parameters must be tuned manually on the real system to achieve the best performance. To avoid this manual tuning, methods from machine learning, such as Bayesian optimization, have been used. However, as these methods evaluate different controller parameters, safety-critical system failures may happen. We overcome this problem by applying, for the first time, a recently developed safe optimization algorithm, SAFEOPT, to the problem of automatic controller parameter tuning. Given an initial, low-performance controller, SAFEOPT automatically optimizes the parameters of a control law while guaranteeing system safety and stability. It achieves this by modeling the underlying performance measure as a Gaussian process and only exploring new controller parameters whose performance lies above a safe performance threshold with high probability. Experimental results on a quadrotor vehicle indicate that the proposed method enables fast, automatic, and safe optimization of controller parameters without human intervention.**

I. INTRODUCTION

An extended version of this paper including a link to the associated Python code is found in [1].

Tuning controller parameters is a challenging task, which requires significant domain knowledge and which can be very time consuming. Classical approaches to automate this process, such as the ones in [2] and [3], either rely on model assumptions (e.g., linearity), which may be the very reason why the initial, model-based controller performs poorly, or require gradient approximations, which are difficult to obtain from noisy measurements. Moreover, gradient-based methods are not guaranteed to find the global optimum.

Recently, Bayesian optimization, a method popular in the field of machine learning, has been used to automate the controller optimization process [4], [5], [6]. In Bayesian optimization, the performance function, which maps controller parameters to performance values, is often modeled as a Gaussian process (GP), which guides the sampling process to informative parameter combinations. As a result, the controller that globally maximizes the performance measure can be found within few evaluations on the real system. Another major advantage of the method is that it explicitly models noise in the performance measure evaluations, which results in a more robust procedure compared to non-Bayesian methods. Moreover, in [7] it was experimentally shown that the Bayesian optimization algorithm in [8] outperforms other Bayesian and non-Bayesian global optimization methods.

Despite experimental success, Bayesian Optimization has one weakness when it comes to real-world experiments. While gradient-ascent methods, such as [3], typically improve at every iteration and thereby ensure that the resulting controllers continue to be stable, informative samples in a Bayesian optimization setting are typically far away from the original control law to gain maximum information. This often leads to the evaluation of unstable controllers and system failures early on in the optimization process.

In this paper, we overcome this problem by using SAFEOPT [9], a Bayesian optimization algorithm that builds on the results from [8] and, in addition, guarantees safety by only evaluating controllers that have a performance above a safe threshold with high probability. The result is a safe, automatic controller tuning algorithm, which we demonstrate in an aerial-vehicle experiment. A video can be found at http://tiny.cc/iros15_video.

II. PROBLEM STATEMENT

The goal of this work is to automatically find the optimal controller parameters for a nonlinear control law, which maximize a given performance measure. The control law may have internal states (e.g., an integrator component). We assume that the overall system is safety-critical; that is, the optimization algorithm must ensure stability when evaluating new controller parameters. In order to start the optimization procedure, we assume an initial set of stabilizing controller parameters (with poor performance) is available.

We encode the safety criterion as a performance threshold below which we do not want to fall with high probability. For example, we may set the threshold at 95% of the performance of the initial control law. Conceptually, this ensures stability, since unstable systems have a significantly lower performance.

III. METHODOLOGY

Our approach builds upon the safe optimization algorithm SAFEOPT [9]. This algorithm models the nonlinear performance function as a GP, where the controller parameters are the inputs and the associated controller performance is the output data. The GP provides not only a mean estimate of the performance function but also corresponding uncertainty information. This information is used to provide high-probability safety guarantees by only evaluating control laws on the real system, where the 3σ (99%) confidence

Felix Berkenkamp and Andreas Krause are with the Learning & Adaptive Systems Group (LAS), ETH Zurich, Switzerland. Email: {befelix, krausea}@ethz.ch

Angela P. Schoellig is with the University of Toronto Institute for Aerospace Studies (UTIAS), Canada. Email: schoellig@utias.utoronto.ca

This research was supported in part by SNSF grant 200020_159557, NSERC grant RGPIN-2014-04634, and the Connaught New Researcher Award.

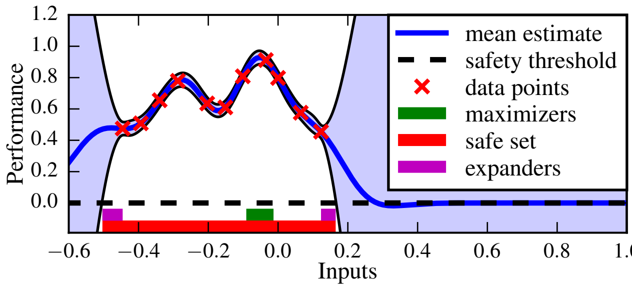

Fig. 1. Based on the mean estimate (blue) and the 3σ confidence interval (light blue), the SAFEOPT algorithm selects safe points (red), which are either potential maximizers (green) or expand the safe set (magenta).

interval of the GP's estimated performance is above a safety threshold.

In this setting, the challenge is to find an evaluation strategy that increases the set of controllers known to be safe (exploration) and simultaneously finds the global maximum within this safe set (exploitation). SAFEOPT provides a solution to this problem by choosing to evaluate the safe controller parameters about whose performance we are most uncertain from two sets: the set of potential maximizers, whose performance may lie above the current maximum according to our GP estimate, and the set of potential expanders, which are parameters that can expand the set of safe controllers (see Fig. 1). For details refer to [9], [1].

IV. RESULTS

We have demonstrated our algorithm on a quadrotor vehicle, the Parrot AR.Drone 2.0, by learning the optimal controller gains for the position controller in x-direction. The position controller generates a pitch command u, which is, in turn, the input to an unknown, on-board attitude controller. The goal is to find the best state-feedback controller $u = \mathbf{K}\mathbf{x}$, with $\mathbf{x} = (x - x_r, \dot{x})$ and $\mathbf{K} = (k_1, k_2)$, which minimizes the cost during a 1-meter reference position change, x_r. Specifically, the cost is $\sum_{k=0}^{N} \mathbf{x}_k^{\mathrm{T}} \mathbf{Q} \mathbf{x}_k + R u_k^2$, with weighting matrices $\mathbf{Q}$ and R, over a time horizon of 5 s ($N = 350$). We define the performance function as the cost improvement relative to 95% of the cost of the initial controller. The safe threshold is set to 0.

We discretize the parameter space uniformly into $10,000$ parameter combinations in $[-0.6, 0.1]^2$, explicitly including positive controller gains, which certainly lead to crashes. We set the initial parameters to $(-0.4, -0.4)$, which lead to poor performance. Lower controller gains lead to instability.

The estimated performance function after 30 experiments is shown in Fig. 2. The optimization routine can be roughly separated into three stages. Initially, the algorithm evaluates parameters close to the initial controller parameters to gain information about the safe set. Once a region of safe controller parameters is determined, the algorithm evaluates the performance function more coarsely in order to expand the safe set. Eventually, the controller is refined by evaluating high-performance parameters that are potential maximizers.

Ultimately, the algorithm identifies the controller gains that maximize the performance measure. Because we omitted the on-board controller and its internal states and due to the nonlinearity of the quadrotor dynamics, the resulting performance function in Fig. 2 is similar to, but not the same as, the quadratic function that one would have expected from linear quadratic control theory.

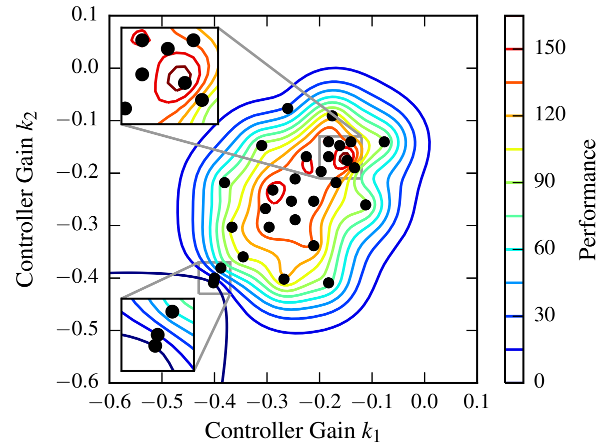

Fig. 2. SAFEOPT adaptively decides where to sample based on safety and informativeness. The bottom-left corner shows the zoomed-in section of the first three samples, which are close together to determine the location of the safe region. The top-left corner shows the zoomed-in section around the maximum, which has more samples to determine the precise location of the maximum. Other areas are more coarsely sampled to expand the safe region.

V. CONCLUSION

We presented the first application of SAFEOPT on a real robotic system by successfully optimizing the position controller of a quadrotor vehicle. It was shown that the algorithm enables efficient, automatic, and global optimization of the controller parameters without risking dangerous

REFERENCES

[1] F. Berkenkamp, A. P. Schoellig, and A. Krause, "Safe controller optimization for quadrotors with Gaussian processes," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2016, (submitted), arXiv:1509.01066 [cs.RO].

[2] K. J. Åström, T. Hägglund, C. C. Hang, and W. K. Ho, "Automatic tuning and adaptation for PID controllers - a survey," *Control Engineering Practice*, vol. 1, no. 4, pp. 699–714, 1993.

[3] N. J. Killingsworth and M. Krstić, "PID tuning using extremum seeking: online, model-free performance optimization," *IEEE Control Systems*, vol. 26, no. 1, pp. 70–79, 2006.

[4] R. Calandra, N. Gopalan, A. Seyfarth, J. Peters, and M. P. Deisenroth, "Bayesian gait optimization for bipedal locomotion," in *Learning and Intelligent Optimization*. Springer, 2014, pp. 274–290.

[5] D. J. Lizotte, T. Wang, M. H. Bowling, and D. Schuurmans, "Automatic gait optimization with Gaussian process regression." in *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)*, vol. 7, 2007, pp. 944–949.

[6] M. Tesch, J. Schneider, and H. Choset, "Using response surfaces and expected improvement to optimize snake robot gait parameters," in *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2011, pp. 1069–1074.

[7] R. Calandra, A. Seyfarth, J. Peters, and M. P. Deisenroth, "An experimental comparison of Bayesian optimization for bipedal locomotion," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 1951–1958.

[8] N. Srinivas, A. Krause, S. M. Kakade, and M. Seeger, "Gaussian process optimization in the bandit setting: no regret and experimental design," in *Proc. of the International Conference on Machine Learning (ICML)*, 2010.

[9] Y. Sui, A. Gotovos, J. W. Burdick, and A. Krause, "Safe exploration for optimization with Gaussian processes," in *Proc. of the International Conference on Machine Learning (ICML)*, 2015, pp. 997–1005.