$\mathcal{L}_1$-$\mathcal{GP}$: $\mathcal{L}_1$ Adaptive Control with Bayesian Learning

Aditya Gahlawat[*]   GAHLAWAT@ILLINOIS.EDU
Pan Zhao[*]   PANZHAO2@ILLINOIS.EDU
Andrew Patterson[*]   APPATTE2@ILLINOIS.EDU
Naira Hovakimyan[*]   NHOVAKIM@ILLINOIS.EDU
Evangelos A. Theodorou[†]   EVANGELOS.THEODOROU@AE.GATECH.EDU

Abstract

We present $\mathcal{L}_1$-$\mathcal{GP}$, an architecture based on $\mathcal{L}_1$ adaptive control and Gaussian Process Regression (GPR) for *safe simultaneous control and learning*. On one hand, the $\mathcal{L}_1$ adaptive control provides stability and transient performance guarantees, which allows for GPR to efficiently and safely learn the uncertain dynamics. On the other hand, the learned dynamics can be conveniently incorporated into the $\mathcal{L}_1$ control architecture without sacrificing robustness and tracking performance. Subsequently, the learned dynamics can lead to less conservative designs for performance/robustness tradeoff. We illustrate the efficacy of the proposed architecture via numerical simulations.

Keywords: Bayesian Learning, Gaussian Process Regression, Safe Adaptive Control

1. Introduction

The historical premise of adaptive control was to control uncertain systems while simultaneously learning the system parameters and providing robustness to uncertainties. Rudolf Kalman was the first to coin the term "self-tuning controller" in 1958 by introducing optimal linear-quadratic regulator (LQR) with explicit identification of parameters (Kalman, 1958). The field of adaptive control since then witnessed tremendous developments, capturing different classes of nonlinear systems, including presence of unmodeled dynamics, switching models, hybrid systems and other singularities, e.g. Åström and Wittenmark (2008); Landau (1979); Narendra et al. (1980); Sastry and Bodson (2011); Ioannou and Sun (2012), and references therein. The main architectures were inspired by inverse Lyapunov design, ensuring asymptotic stability in the presence of system uncertainties and disturbances. Recent developments in $\mathcal{L}_1$ adaptive control filled the last gap of explicitly introducing robustness into the problem formulation, leading to a framework with *a priori* guaranteed robustness margins, transient and steady-state specifications (Cao and Hovakimyan, 2008; Hovakimyan and Cao, 2010). In $\mathcal{L}_1$ control architecture, estimation is decoupled from control, thereby allowing for arbitrarily fast adaptation subject only to hardware limitations. The $\mathcal{L}_1$ control has been successfully implemented on NASA's AirStar 5.5% subscale generic transport aircraft model (Gregory et al., 2009, 2010) and Calspan's Learjet (Ackerman et al., 2016, 2017) and F16 aircraft and unmmaned aerial vehicles (Kaminer et al., 2010, 2015; Jafarnejadsani et al., 2017; Zuo and Ru, 2014). Despite these vast developments, the issue of learning the system dynamics and/or uncertainties remained unresolved, as the typical estimation schemes in all these adaptive architectures require persistency of excitation (PE) type assumption on reference signals to ensure parameter convergence. Such requirement is unacceptable in safety-critical applications, rendering the conventional Lyapunov-based adaptive control architectures incomplete, if parameter/system identification is to be addressed *simultaneously* with transient specifications.

[*] Mechanical Science and Engineering, University of Illinois at Urbana Champaign, Urbana, IL-61801
[†] Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA-30332

The last two decades have witnessed a type of data explosion that has revolutionized the industry of autonomous systems. Tools from machine learning have been extensively explored in modeling, identification, and control of dynamic systems. A few examples of such tools include, but are not limited to, neural networks (Lewis et al., 1998), Gaussian processes (Williams and Rasmussen, 2006), and reinforcement learning (Sutton and Barto, 2018). In many of these instances, guarantees of stability have not been prioritized, yet having an impressive demonstration was the main objective to show the power of data-driven methods towards achieving full autonomy (Lillicrap et al., 2015; Deisenroth et al., 2013; Levine et al., 2016; Pan et al., 2020). Due to its data efficiency, the nonparameteric structure and the ability to provide uncertainty quantification, Gaussian Process Regression (GPR) has become popular in safety-critical learning and control (Aswani et al., 2013; Akametalu et al., 2014; Berkenkamp and Schoellig, 2015; Berkenkamp et al., 2017; Hewing et al., 2019; Wang et al., 2018), including application to model reference adaptive control (Chowdhary et al., 2014). When the learning methods generate unsafe reference/control commands, the control barrier function methods presented by Cheng et al. (2019), Salehi et al. (2019), and Taylor and Ames (2019) correct the control input to ensure the system state remains in a safe set. In the present work the desired trajectory is designed to be feasible and safe for an appropriately designed reference system. The safety and feasibility guarantees are then dependent on the ability of an adaptive controller to emulate the reference system. This design philosophy allows safe and feasible trajectories to be generated *a priori*, instead of relying on run-time optimization routines to correct the unsafe trajectories. In most of the learning control methods presented, the control performance is a direct function of the quality of the learned uncertainties. In this paper we combine the formal stability and robustness guarantees of $\mathcal{L}_1$ adaptive control with Gaussian Processes to ensure safe learning and adaptation with *a priori* transient bounds. This would enable the satisfaction of control objectives like trajectory tracking and simultaneously enable learning from the collected data.

Over the last two years $\mathcal{L}_1$ control has been explored within NASA's Learn-To-Fly (L2F) framework, wherein a real-time system identification toolbox is integrated across the flight envelope to continuously update the model parameters and enable autonomous flight without wind-tunnel testing, while an $\mathcal{L}_1$ adaptive controller is used to provide robustness and stability guarantees (Snyder, 2019). Incorporation of learning via neural network in $\mathcal{L}_1$ control was investigated in Cooper et al. (2014). The system identification within L2F and the neural network based learning require some prior knowledge of the system and uncertainty structure to facilitate parameter estimation.

In this paper we explore the $\mathcal{L}_1$ control architecture with Bayesian learning in the form of GPR for safe learning with guaranteed stability and control performance throughout the learning phase. We assume no availability of model structure and resort to the GPR to learn the uncertain dynamics whenever possible, while achieving given control objectives like trajectory tracking. The predictor in $\mathcal{L}_1$ adaptive control architecture naturally allows the incorporation of the availabe knowledge in a systematic way[1]. We demonstrate that one can learn model uncertainties efficiently and safely via GPR, while guaranteeing the stability and performance. Furthermore, we illustrate that the fast adaptation of $\mathcal{L}_1$ controller intervenes when the uncertainties change. This ensures safe control while the Bayesian learning catches up.

Finally, one may argue that if $\mathcal{L}_1$ adaptive control already guarantees stability and robustness, then why incorporate learning within it. Instead, the learning should be kept separate if the goal is just safe learning. While this assertion is true, in addition to safe learning, we are also demonstrating

---

1. The *a priori* knowledge of a system such as time-delay and input saturation can be conveniently incorporated into the state predictor, which helps to improve both the performance and robustness (Kharisov et al., 2011).

that learning can be incorporated within the $\mathcal{L}_1$ architecture without harming robustness or performance. This is the initial step of the envisioned research, where the next step is to illustrate how the learning can improve performance, without sacrificing robustness, when a larger operational envelope is considered as compared to a single trim condition. On the other hand, the benefits of $\mathcal{L}_1$-$\mathcal{GP}$ for purposes of planning (guidance and navigation) in highly uncertain environments are yet to be illustrated on appropriate benchmark examples.

The problem formulation is introduced in Section 2, where an overview of GPR and $\mathcal{L}_1$ adaptive control is also provided. The main architecture of $\mathcal{L}_1 - \mathcal{GP}$ is presented in Section 3. Numerical validation is provided in Section 4. The manuscript is concluded in Section 5.

2. Problem Formulation

Let $\| \cdot \|_p$ denote the p-norm defined on the space $\mathbb{R}^n$, $n \in \mathbb{N}$, and $\| \cdot \|$ denote the 2-norm. $\mathbb{I}_n$ denotes an identity matrix of size n. Given a positive scalar κ, we denote by $\mathbb{X}_\kappa$ the compact set containing all $x \in \mathbb{R}^n$ such that $\|x\|_\infty \leq \kappa$. Similarly, arbitrary compact subsets of $\mathbb{R}^n$ are denoted by $\mathbb{X}$. For any time-varying function $g(t)$, $g(s)$ denotes its Laplace transform when it exists, and $\|g\|_{\mathcal{L}_\infty}$ denotes its $\mathcal{L}_\infty$ norm. For a transfer function matrix $G(s)$, $\|G(s)\|_{\mathcal{L}_1}$ denotes its $\mathcal{L}_1$-norm. Next we discuss the problem formulation by considering the following system:

$$\dot{x}(t) = A_m x(t) + B_m(u(t) + f(x(t))), \quad x(0) = x_0, \quad \text{and} \quad y(t) = C_m x(t), \qquad (1)$$

where $x(t) \in \mathbb{R}^n$ is the system state, $u(t) \in \mathbb{R}^m$ is the control input, $A_m \in \mathbb{R}^{n \times n}$ is a known Hurwitz matrix specifying the desired closed-loop dynamics, $B_m \in \mathbb{R}^{n \times m}$ and $C_m \in \mathbb{R}^{m \times n}$, $m \leq n$, are known matrices with $\mathrm{rank}(B_m) = m$, $f : \mathbb{R}^n \to \mathbb{R}^m$ is the *unknown* nonlinearity representing the model uncertainties, and $y(t) \in \mathbb{R}^m$ is the regulated output. The matrices A_m, B_m and C_m are the reference system matrices specifying the desired closed-loop system behavior.

Assumption 1 *The constituent functions of the unknown nonlinearity $f = \begin{bmatrix} f_i & \cdots & f_m \end{bmatrix}^\top$, $f_i : \mathbb{R}^n \to \mathbb{R}$ are samples from Gaussian processes $\mathcal{GP}(0, K_{f,i}(x, x'))$, where the kernels $K_{f,i} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and their Lipschitz constants $L_{k,i}(\mathbb{X})$ on compact subsets of $\mathbb{R}^n$ are known.*

Assumption 2 *There exists a known conservative bound $L_f(\mathbb{X})$ such that $\|\nabla_x f(x)\|_\infty \leq L_f(\mathbb{X})$ for all $x \in \mathbb{X}$, and B_0 such that $\|f(0)\|_\infty \leq B_0$.*

The objective is to learn the model uncertainty f and track a given bounded reference signal $r(t)$ with quantifiable performance bounds both in transient and steady-state. Next we discuss the two components of our approach, namely GPR and $\mathcal{L}_1$ adaptive control.

2.1. Bayesian Learning of Model Uncertainties

We present the high-probability bounds for the uniform prediction errors by first setting up the measurement model. Assume we have $N \in \mathbb{N}$ measurements of the form $y_j = f(x_j) + \zeta = \left(B_m^\top B_m\right)^{-1} B_m^\top (\dot{x}_j - A_m x_j) - u_j + \zeta \mathbb{I}_m$, $\zeta \sim \mathcal{N}(0, \sigma_n^2)$, $y_j \in \mathbb{R}^m$. Note that, estimates of $\dot{x}$ may be numerically generated with the estimation errors incorporated into ζ, for example, by the Savitsky-Golay filter (Schafer, 2011). We define the data set as $\mathcal{D}_N = \{\mathbf{Y}, \mathbf{X}\}$, where $\mathbf{Y} \in \mathbb{R}^{N \times m}$, $\mathbf{X} \in \mathbb{R}^{N \times n}$ and are defined as $\mathbf{Y} = \begin{bmatrix} y_1 & \cdots & y_N \end{bmatrix}^\top$, and $\mathbf{X} = \begin{bmatrix} x_1 & \cdots & x_N \end{bmatrix}^\top$. GPR proceeds by using the assumption that $f_i \sim \mathcal{N}(0, K_{f_i}(x, x'))$, $i \in \{1, \ldots, m\}$, and the data $y_j \sim \mathcal{N}(f(x_j), \sigma_n^2 \mathbb{I}_m)$ to formulate the posterior distributions conditioned on data at any test point $x^\star \in \mathbb{R}^n$ as

$$f_i(x^\star)|\mathbf{Y}_i \sim \mathcal{N}(\mu_i(x^\star), \sigma_i^2(x^\star)), \quad i \in \{1, \ldots, m\}, \qquad (2)$$

where $\mathbf{Y}_i$ is the ith column of $\mathbf{Y}$. The terms $\mu_i(x^\star)$ and $\sigma_i(x^\star)$ are mean and variance of the GP model and are defined as $\mu_i(x^\star) = \mathbf{K}_i^\star(x^\star)^\top \left(\mathbf{K}_i + \sigma_n^2\mathbb{I}_N\right)^{-1}\mathbf{Y}_i$, and $\sigma_i^2(x^\star) = \mathbf{K}_i^{\star\star}(x^\star) - \mathbf{K}_i^\star(x^\star)^\top \left(\mathbf{K}_i + \sigma_n^2\mathbb{I}_N\right)^{-1}\mathbf{K}_i^\star(x^\star)$. The terms $\mathbf{K}_i^{\star\star}(x^\star)$, $\mathbf{K}_i^\star(x^\star)$ and $\mathbf{K}_i$ are defined based on the kernel of GP model as $\mathbf{K}_i^{\star\star}(x^\star) = K_{f,i}(x^\star, x^\star) \in \mathbb{R}$, $\mathbf{K}_i^\star(x^\star) = K_{f,i}(\mathbf{X}, x^\star) \in \mathbb{R}^N$, $\mathbf{K}_i = K_{f,i}(\mathbf{X}, \mathbf{X}) \in \mathbb{R}^{N \times N}$. Further details can be found in Williams and Rasmussen (2006) and Bishop (2006). A major advantage of GPR is that the predictive estimates are in the form of predictive distributions, as in (2), as opposed to point estimates. These predictive distributions can be used to produce high probability bounds on the prediction errors as in Srinivas et al. (2012); Chowdhury and Gopalan (2017). Recently, Lederer et al. (2019) presented a method of computing such bounds, which are amenable to on-line computation, the generalization of which is presented below.

Theorem 1 *Let Assumptions 1- 2 hold. Given the distributions in (2), for some $\xi > 0$ and any compact $\mathbb{X} \subset \mathbb{R}^n$, let $\mu(x) = \begin{bmatrix} \mu_1(x) & \cdots & \mu_m(x) \end{bmatrix}^\top$, $\sigma(x) = \begin{bmatrix} \sigma_1(x) & \cdots & \sigma_m(x) \end{bmatrix}^\top$, and*

$$L_{\mu_i}(\mathbb{X}) = L_{k,i}(\mathbb{X})\sqrt{N}\|(\mathbf{K}_i + \sigma_n^2\mathbb{I}_N)^{-1}\mathbf{Y}_i\|, \quad L_\mu(\mathbb{X}) = \max_{i \in \{1,\dots,m\}} L_{\mu_i}(\mathbb{X})$$

$$\omega_{\sigma_i}(\xi) = \sqrt{2\xi L_{k,i}(\mathbb{X})\left(1 + N\|(\mathbf{K}_i + \sigma_n^2\mathbb{I}_N)^{-1}\| \max_{x,x' \in \mathbb{X}} K_{f,i}(x, x')\right)}, \quad \omega_\sigma(\xi) = \max_{i \in \{1,\dots,m\}} \omega_{\sigma_i}(\xi),$$

for $i \in \{1, \dots, m\}$. Furthermore, for any $\delta \in (0,1)$ define $\beta(\xi) = 2\log\left(\frac{mM(\xi,\mathbb{X})}{\delta}\right)$, $\gamma(\xi) = \left(\frac{L_f(\mathbb{X})}{n} + L_\mu(\mathbb{X})\right)\xi + \sqrt{\beta(\xi)}\omega_\sigma(\xi)$, where $M(\xi,\mathbb{X})$ is the ξ-covering number of $\mathbb{X}$. Then, we have

$$Pr\left\{\|f(x) - \mu(x)\|_\infty \le e_f(x) = \sqrt{\beta(\xi)}\|\sigma(x)\|_\infty + \gamma(\xi), \quad \forall x \in \mathbb{X}\right\} \ge 1 - \delta.$$

The proof is provided in the extended version of the manuscript in Gahlawat et al. (2020).

2.2. Overview of $\mathcal{L}_1$ Adaptive Control

We now briefly review the $\mathcal{L}_1$ control architecture for (1) *without* incorporation of learned dynamics. Further details can be found in Hovakimyan and Cao (2010). An $\mathcal{L}_1$ controller consists of a state predictor, an adaptation law, and a control law. The **state predictor** is given as

$$\dot{\hat{x}}(t) = A_m\hat{x}(t) + B_m(u(t) + \hat{\sigma}(t)), \; \hat{x}(0) = \hat{x}_0, \quad \text{and} \quad \hat{y}(t) = C_m\hat{x}(t), \tag{3}$$

where $\hat{x}(t) \in \mathbb{R}^n$ is the predictor state and $\hat{x}_0$ is its initial value (that may be different from x_0 in (1)), and $\hat{\sigma}(t) \in \mathbb{R}^m$ is the adaptive estimate. The **adaptive estimate** is updated according to

$$\hat{\sigma}(t) = \hat{\sigma}(iT_s), \quad \hat{\sigma}(iT_s) = -B_m^+\Phi^{-1}(T_s)e^{A_mT_s}\tilde{x}(iT_s), \tag{4}$$

where $t \in [iT_s, (i+1)T_s]$ with T_s being the sampling time and $i \in \mathbb{N}$, B_m^+ is the pseudo-inverse of B_m, $\Phi(T_s) \triangleq A_m^{-1}(e^{A_mT_s} - \mathbb{I}_n)$, and $\tilde{x}(t) \triangleq \hat{x}(t) - x(t)$. The **control law** is given as

$$u(s) = C(s)(\hat{\sigma}(s) - k_g r(s)), \tag{5}$$

where $\hat{\sigma}(s)$ is the Laplace transform of $\hat{\sigma}(t)$, $r(t)$ is the reference signal and $k_g \triangleq -(C_mA_m^{-1}B_m)^{-1}$ is a feedforward gain to ensure that the desired transfer function matrix $M(s) = C_m(s\mathbb{I}_n - A_m)^{-1}B_m$ has an identity DC gain, and $C(s)$ is a lowpass filter with $C(0) = \mathbb{I}_m$, subject to

$$\|H(s)(\mathbb{I} - C(s))\|_{\mathcal{L}_1} < \left(\rho_r - \|H(s)C(s)k_g\|_{\mathcal{L}_1}\|r\|_{\mathcal{L}_\infty} - \rho_{\text{in}}\right) / (L_f(\mathbb{X}_{\rho_r})\rho_r + B_0), \tag{6}$$

where $H(s) \triangleq (s\mathbb{I}_n - A_m)^{-1}B_m$, $\rho_{\text{in}} \triangleq \left\| s(s\mathbb{I} - A_m)^{-1} \right\|_{\mathcal{L}_1} \rho_0$ with ρ_0 being a known bound for the initial state x_0 (i.e. $\|x_0\|_\infty \leq \rho_0$), B_0 and $L_f(\cdot)$ are defined in Assumption 2, ρ_r is a positive constant that defines the semiglobal domain of attraction. The filter can be designed via optimization (Hovakimyan and Cao, 2010, Section 2.6), (Jafarnejadsani et al., 2017). When there is no initialization error, i.e. $\hat{x}_0 = x_0$, following Hovakimyan and Cao (2010), if $T_s \to 0$, then the state and control signals of the closed-loop $\mathcal{L}_1$ system – both in transient and steady-state – can be made arbitrarily close to the corresponding signals of the following non-adaptive auxiliary reference system

$$\dot{x}_{\text{ref}}(t) = A_m x_{\text{ref}}(t) + B_m(u_{\text{ref}}(t) + f(x_{\text{ref}}(t))), \quad x_{\text{ref}}(0) = x_0, \tag{7a}$$

$$u_{\text{ref}}(s) = C(s)(k_g r(s) - \eta_{\text{ref}}(s)), \quad y_{\text{ref}}(t) = C_m x_{\text{ref}}(t), \tag{7b}$$

where $\eta_{\text{ref}}(s)$ is the Laplace transform of $\eta_{\text{ref}}(t) \triangleq f(x_{\text{ref}}(t))$. In the presence of non-zero initialization error, the performance bounds between the adaptive system and the reference system will contain additive exponentially decaying terms that depend on the initialization error. The reference system defines the *ideal achievable performance*, where the uncertainty is perfectly known and cancelled within the bandwidth of the filter $C(s)$. Its stability hinges upon the same condition in (6), while the bandwidth of the filter $C(s)$ defines the tradeoff between performance and robustness.

3. The $\mathcal{L}_1$-$\mathcal{GP}$ Architecture

The architecture of the $\mathcal{L}_1$-$\mathcal{GP}$ controller contains two primary components: i) the Bayesian learner that uses a GPR algorithm to produce estimates of the uncertainty f, and ii) the $\mathcal{L}_1$ adaptive controller which incorporates the estimates and generates the control input $u(t)$.

Bayesian learner: The task of the Bayesian learner is to use the collected data to produce the estimates of the uncertainty f in the form of the mean function μ of the posterior distribution. Furthermore, it also outputs the high-probability prediction error bounds presented in Theorem 1 as

$$\mathcal{M}(x(t), t) = \{\hat{f}(x(t), t), \hat{e}_f(x(t), t)\}, \tag{8}$$

where the piecewise static in time $\hat{f}$ and $\hat{e}$ are defined as $\hat{f}(x(t), t) = \mu_k(x(t))$ and $\hat{e}_f(x(t), t) = e_{f,k}(x(t))$, for all $t \in [t_k, t_{k+1})$, $t_k \in \mathcal{T}$. Here, $\mathcal{T}$ is the set of discrete time-instances at which the Bayesian learner updates the model parameters. Thus, over the time interval $[t_k, t_{k+1})$, $\mu_k(x(t)) = \begin{bmatrix} \mu_{k,1}(x(t)) & \cdots & \mu_{k,m}(x(t)) \end{bmatrix}$, where $\mu_{k,i}(\cdot)$ are the mean functions obtained after the kth-model update computed via the posterior distributions in (2). Similarly, $e_{f,k}(x(t))$ is the uniform error bound computed via Theorem 1 after the kth model update. The Bayesian learner updates the model once $N \in \mathbb{N}$ new data points have been collected; thus N is a design parameter. The Bayesian learner is initialized to $\mu_0(x(t)) = 0_m$, which is the prior mean, and $e_{f_0}(x(t)) = e_f(x(t))$ is obtained based solely on the GP priors on f.

Incorporating Learning into $\mathcal{L}_1$ Control: Next, we present the $\mathcal{L}_1$-$\mathcal{GP}$ controller that incorporates the model updates produced by the Bayesian learner into the $\mathcal{L}_1$ controller. Same as the $\mathcal{L}_1$ controller, the $\mathcal{L}_1$-$\mathcal{GP}$ controller consists of the state-predictor, adaptation law, and the control law. The $\mathcal{L}_1$-$\mathcal{GP}$ state predictor is given by

$$\dot{\hat{x}}(t) = A_m \hat{x}(t) + B_m(f_L(t) + \hat{\sigma}(t) + u(t)), \quad \hat{x}(0) = \hat{x}_0 \text{ and } \hat{y}(t) = C_m \hat{x}(t), \tag{9}$$

where $\hat{\sigma}(t)$ is the adaptive estimate of uncertainties, $f_L(t)$ is the solution of the following equation

$$\dot{f}_L(t) = -\omega(t)\left(f_L(t) - \hat{f}(x(t), t)\right), \quad f_L(0) = 0, \tag{10}$$

(*a*) State evolution
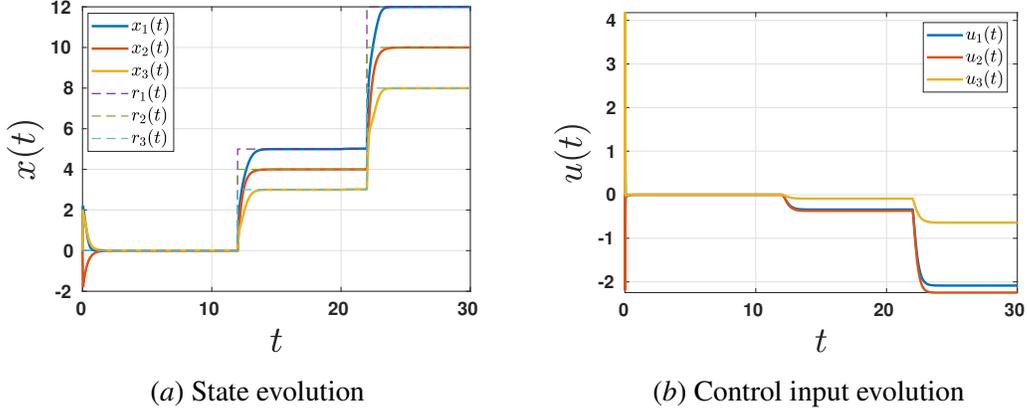
(*b*) Control input evolution

Figure 1: State and control input evolution for $\mathcal{L}_1$-$\mathcal{GP}$ closed-loop system for step reference inputs.

with $\hat{f}(x(t), t)$ being defined in (8), and

$$\omega(s) = L(s)\hat{\omega}(s), \quad \hat{\omega}(t) = \min\{\omega_0/\hat{e}_f(x(t), t), \omega_c\}. \tag{11}$$

Here, $\omega_0 > 0$ is an an arbitrarily small *a priori* chosen constant, and ω_c is the bandwidth of $C(s)$ verifying the $\mathcal{L}_1$-norm condition in (6), $\hat{e}_f(x(t), t)$ is the output of the Bayesian learner in (8), and $L(s)$ is a low-pass filter. The update of the adaptive estimate $\hat{\sigma}$ is governed by the piecewise-constant adaptation law with sampling time T_s per (4). The $\mathcal{L}_1$-$\mathcal{GP}$ controller is given by

$$u(s) = -f_L(s) - C(s)(\hat{\sigma}(s) - k_g r(s)). \tag{12}$$

Note that $\hat{e}_f(x(t), t)$, defined in (8), starts at $e_{f_0}(x(t))$ when no model updates have been performed, and ideally approaches zero after sufficiently large number of model updates have been performed as the size of the data set increases. Therefore, by the law presented in (11), $\omega(t)$ in (10) increases from an arbitrarily small value $\omega_0/e_{f_0}(x(t))$ to ω_c, the bandwidth of the filter $C(s)$. Moreover, the change in $\omega(t)$ is smooth because of the low-pass filter $L(s)$. The filter (10) allows incorporation of the learned uncertainties smoothly into the system. Further, as $\hat{f} \to f$ [2], it is expected that $\tilde{x}(t)$ and $\hat{\sigma}(t)$ go to zero. Thus, the $\mathcal{L}_1$-$\mathcal{GP}$ closed-loop system (1), (9)-(12) converges to the $\mathcal{L}_1$ reference system in (7). The adaptive estimate $\hat{\sigma}$ is driven by the prediction error $\tilde{x} \triangleq \hat{x} - x$, given by

$$\dot{\tilde{x}}(t) = A_m \tilde{x}(t) + B_m \left(f_L(t) - f(x(t)) + \hat{\sigma}(t) \right), \quad \tilde{x}(0) = \hat{x}_0 - x_0. \tag{13}$$

The learned dynamics are used to cancel the model uncertainty via $f_L(t)$ in (10). From the prediction error dynamics (13), it is evident that the $-C(s)\hat{\sigma}(s)$ component of the control law (12) compensates for the remaining uncertainty, $f(x(t)) - f_L(t)$, within the bandwidth of the filter $C(s)$.

Remark 2 *Proof of the stability of the $\mathcal{L}_1$-$\mathcal{GP}$ closed-loop system can be established by following the ideas in Cooper et al. (2014); Snyder (2019).*

4. Simulation Results

We now present the results of numerical experimentation. We consider the dynamics of body-frame angular rates $x(t) \in \mathbb{R}^3$ of a multirotor craft given by

$$\dot{x}(t) = -J^{-1}\left(x(t) \times Jx(t)\right) + J^{-1}f(x(t)) + J^{-1}u_{total}(t), \quad y(t) = x(t), \tag{14}$$

---

2. The expression $\hat{f} \to f$ implies that the high-probability bounds on $\|f(x) - \mu(x)\|_\infty$ go to zero. The conditions under which this convergence takes place can be found in Lederer et al. (2019).

with $x(0) = x_0 = 0_3$, where $J = \text{diag}\{0.011, 0.011, 0.021\}$ is the known moment-of-inertia matrix, $f(x(t))$ is the model uncertainty, and $u_{total}(t) \in \mathbb{R}^3$ is the control input, which, for a multirotor craft presents the body-frame moments. The control input is decomposed as $u_{total}(t) = u_{bl}(t) + u(t)$, where $u_{bl}(t)$ is the baseline input and $u(t)$ is the $\mathcal{L}_1$-$\mathcal{GP}$ input. The role of the baseline input is to inject desired dynamics, i.e., $u_{bl}(t) = JA_m x(t) + (x(t) \times Jx(t))$, where $A_m = -3\mathbb{I}_3$. With baseline input injected into (14), the partially closed-loop system can be written in the form of (1) with $B_m = J^{-1}$ and $C_m = \mathbb{I}_3$. Next, we consider the following model uncertainty

$$f(x(t)) = \begin{bmatrix} 0.01\left(x_1^2(t) + x_3^2(t)\right) & 0.01\left(x_3(t)x_2(t) + x_1^2(t)\right) & 0.01\left(x_3^2(t)\right) \end{bmatrix}^\top. \qquad (15)$$

For the $\mathcal{L}_1$-$\mathcal{GP}$ control input, we set $C(s) = \omega_c/(\omega_c + s)\mathbb{I}_3$, $\omega_c = 80\ rad/s$, $L(s) = 0.01/(0.01 + s)$, and $\omega_0 = 1$. The predictor (9) is initialized with $\hat{x}_0 = \begin{bmatrix} 0.5 & 0.5 & 0.5 \end{bmatrix}^\top$, which is distinct from the system's initial conditions in (14). For the GPR, we choose the Squared-Exponential (SE) kernels as $K_{f,i}(x, x') = \sigma_f^2 \exp\left(-(x - x')^\top(x - x')/2l^2\right)$, where the unoptimized hyper-parameters are chosen to be $\sigma_f = l = 1$. Furthermore, we upper bound the covering number $\beta(\xi)$ (Thm. 1) as in Lederer et al. (2019) using $\xi = 0.001$ and conservatively chosen $\mathbb{X} = \{x \in \mathbb{R}^3 : \|x\|_\infty \leq 15\}$. For the purposes of simulation, we ignore the $\gamma(\xi)$ term (Thm. 1) as it can be made arbitrarily small. Finally, we choose $\delta = 0.01$, the feedforward

Figure 2: Evolution of $\|f_L(t)\|$ and $\|\eta(t)\|$ for sinusoidal reference commands.

gain $k_g = -\left(C_m A_m^{-1} B_m\right)^{-1}$ and the sampling time for the update of the adaptive estimate $\hat{\sigma}(t)$ as $T_s = 0.001$. The Bayesian learner collects data at the rate $1\ Hz$ and updates the model after $N = 10$ new data-points have been collected; thus the model is updated at $0.1\ Hz$. Figure 1 illustrates the state evolution and the $\mathcal{L}_1$-$\mathcal{GP}$ input u in response to a step reference command. The figure shows the scaled response of the system without retuning, a property that $\mathcal{L}_1$-$\mathcal{GP}$ shares with $\mathcal{L}_1$ control. Moreover, $\mathcal{L}_1$-$\mathcal{GP}$ preserves the performance bounds which are guaranteed for $\mathcal{L}_1$ control. Next we show the effect of learning within the $\mathcal{L}_1$-$\mathcal{GP}$ input $u(t)$. Recall that $u(t)$ in (12) is comprised of two components, the learning based input $f_L(t)$ and the adaptive input $\eta(t)$, where $\eta(s) = C(s)\hat{\sigma}(s)$. The evolution of these individual components for a sinusoidal reference is given in Fig. 2. Note that the dominant component of the input $u(t)$ transitions from adaptive input $\eta(t)$ to the learning based input $f_L(t)$ as the learning improves. We now demonstrate the safe-learning enabled by the $\mathcal{L}_1$-$\mathcal{GP}$ controller under sudden change of uncertainties. As illustrated in Fig. 2, as the learning improves, the learning based component $f_L(t)$ becomes the major contributor to $u(t)$. However, the adaptive component, $\eta(t)$, always remains active in the background ready to intervene when new uncertainties enter the dynamics. This is crucial for stability and performance guarantees as the learning runs on a long time scale, whereas the fast adaptation due to $\hat{\sigma}(t)$ immediately intervenes to compensate for the new uncertainties. To demonstrate this, the $\mathcal{L}_1$-$\mathcal{GP}$ controller is tasked with tracking a sinusoidal reference command. At $t = 35\ s$, we switch the model uncertainty from $f(x(t))$ in (15) to $f(x(t)) = \begin{bmatrix} 0.5\sin(x_1(t)) & 0.01\cos(x_3(t)) & 0.5\left(\sin(x_1(t)) + \cos(x_2(t))\right) \end{bmatrix}^\top$. The results are given in Fig. 3. At $t = 35\ s$, when the uncertainty $f(x(t))$ switches, the adaptive element $\eta(t)$ immediately intervenes to compensate for the new uncertainty. Furthermore, at this point, the previously learned input $f_L(t)$ is incapable of cancelling the new $f(x(t))$. Therefore,
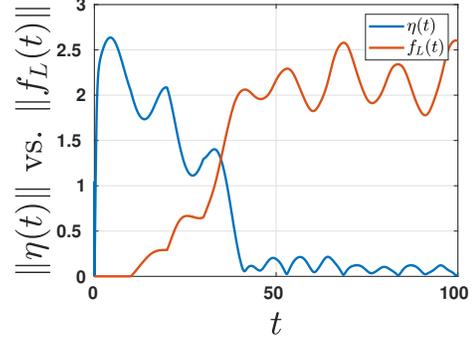
$\eta(t)$ considers $f_L(t)$ as a disturbance to be rejected. However, since $f_L(t)$ enters the system via the low-pass filter (10), it always remains within the bandwidth of $C(s)$, and thus is compensated by the adaptive element $\eta(t)$. Finally, the state evolution illustrates the maintenance of stability of the system. We would also like to remark that both the $\mathcal{L}_1$-$\mathcal{GP}$ and the $\mathcal{L}_1$ control maintain the same
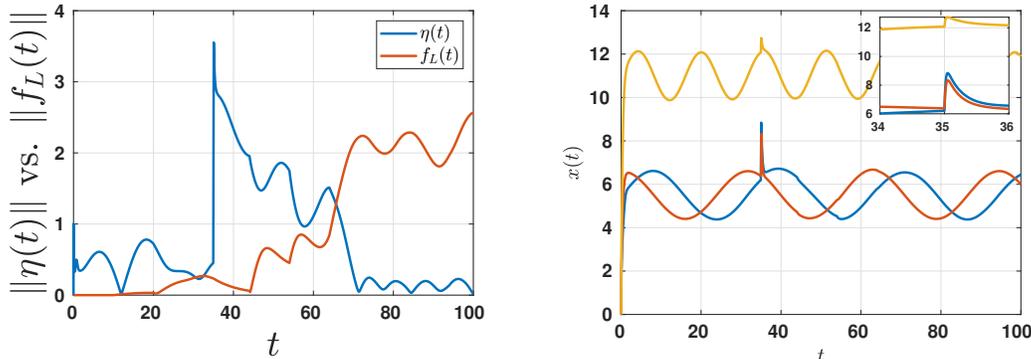
(*a*) Evolution of $\|f_L(t)\|$ and $\|\eta(t)\|$.

(*b*) State evolution. Inset shows the smooth response of the system state across the uncertainty switch.

Figure 3: Learning and adaptive components of the $\mathcal{L}_1$-$\mathcal{GP}$ input $u(t)$ and system state evolution with model uncertainty switch at $t = 35s$.

time-delay margins. The time-delay margins for both control schemes were computed numerically to be $\approx 20\ ms$. This is not surprising since the time-delay margins are dominated by the adaptive elements including the low-pass filter $C(s)$ and sampling time T_s, which are the same for the $\mathcal{L}_1$-$\mathcal{GP}$ and the $\mathcal{L}_1$ controllers.

5. Conclusion

We presented the $\mathcal{L}_1$-$\mathcal{GP}$ architecture, which incorporates Bayesian learning via Gaussian Process Regression (GPR) into the $\mathcal{L}_1$ adaptive control framework. Within the framework, GPR allows for sample-efficient learning of the model uncertainties, while the $\mathcal{L}_1$ controller provides stability, robustness and performance guarantees throughout the learning phase. The $\mathcal{L}_1$-$\mathcal{GP}$ architecture is the initial phase of the research and will next proceed by using learning to improve the performance over a larger envelope of operation, while maintaining given robustness specifications. Eventually, the presented work will be extended to safe and robust planning and control of uncertain systems. The $\mathcal{L}_1$-$\mathcal{GP}$ architecture will be extended to consider spatio-temporal learning for realistic scenarios as most real systems are subject to time-varying disturbances. Further extensions of the $\mathcal{L}_1$-$\mathcal{GP}$ architecture to the case of output-feedback and stochastic systems will also be investigated.

Acknowledgements

References

Kasey A. Ackerman, Enric Xargay, Ronald Choe, Naira Hovakimyan, M. Christopher Cotting, Robert B. Jeffrey, Margaret P. Blackstun, T. Paul Fulkerson, Timothy R. Lau, and Shawn S. Stephens. $\mathcal{L}_1$ stability augmentation system for Calspan's variable-stability Learjet. In *Proceedings of AIAA Guidance, Navigation and Control Conference*, San Diego, California, USA, January 2016. AIAA 2016-0631.

Kasey A. Ackerman, Enric Xargay, Ronald Choe, Naira Hovakimyan, M. Christopher Cotting, Robert B. Jeffrey, Margaret P. Blackstun, T. Paul Fulkerson, Timothy R. Lau, and Shawn S. Stephens. Evaluation of an $\mathcal{L}_1$ adaptive flight control law on Calspan's variable-stability Learjet. *AIAA Journal of Guidance, Control, and Dynamics*, 40(4):1051–1060, 2017.

Anayo K. Akametalu, Jaime F. Fisac, Jeremy H. Gillula, Shahab Kaynama, Melanie N. Zeilinger, and Claire J. Tomlin. Reachability-based safe learning with Gaussian processes. In *Proceedings of 53rd IEEE Conference on Decision and Control*, pages 1424–1431, Los Angeles, CA, USA, 2014.

Karl J. Åström and Björn Wittenmark. *Adaptive Control*. Dover, New York, 2nd edition, 2008.

Anil Aswani, Humberto Gonzalez, S. Shankar Sastry, and Claire Tomlin. Provably safe and robust learning-based model predictive control. *Automatica*, 49(5):1216–1226, 2013.

Felix Berkenkamp and Angela P. Schoellig. Safe and robust learning control with Gaussian processes. In *Proceedings of 2015 European Control Conference (ECC)*, pages 2496–2501, Linz, Austria, 2015.

Felix Berkenkamp, Matteo Turchetta, Angela P. Schoellig, and Andreas Krause. Safe model-based reinforcement learning with stability guarantees. In *Proceedings of 31st International Conference on Neural Information Processing Systems*, pages 908–918, 2017.

Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

Chengyu Cao and Naira Hovakimyan. Design and analysis of a novel $\mathcal{L}_1$ adaptive control architecture with guaranteed transient performance. *IEEE Transactions on Automatic Control*, 53(2): 586–591, 2008.

Richard Cheng, Gábor Orosz, Richard M. Murray, and Joel W. Burdick. End-to-end safe reinforcement learning through barrier functions for safety-critical continuous control tasks. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*, pages 3387–3395, Honolulu, Hawaii, USA, 2019.

Girish Chowdhary, Hassan A. Kingravi, Jonathan P. How, and Patricio A. Vela. Bayesian nonparametric adaptive control using Gaussian processes. *IEEE Transactions on Neural Networks and Learning Systems*, 26(3):537–550, 2014.

Sayak Ray Chowdhury and Aditya Gopalan. On kernelized multi-armed bandits. In *Proceedings of the 34th International Conference on Machine Learning*, pages 844–853, Sydney, Australia, 2017.

John Cooper, Jiaxing Che, and Chengyu Cao. The use of learning in fast adaptation algorithms. *International Journal of Adaptive Control and Signal Processing*, 28(3-5):325–340, 2014.

Marc Peter Deisenroth, Dieter Fox, and Carl Edward Rasmussen. Gaussian processes for data-efficient learning in robotics and control. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(2):408–423, 2013.

Aditya Gahlawat, Pan Zhao, Andrew Patterson, Naira Hovakimyan, and Evangelos A. Theodorou. $\mathcal{L}_1$-$\mathcal{GP}$: $\mathcal{L}_1$ adaptive control with bayesian learning. *arXiv preprint arXiv:2004.14594 [eess.SY]*, 2020.

Irene Gregory, Chengyu Cao, Enric Xargay, Naira Hovakimyan, and Xiaotian Zou. $\mathcal{L}_1$ adaptive control design for NASA AirSTAR flight test vehicle. In *Proceedings of AIAA Guidance, Navigation, and Control Conference*, Chicago, Illinois, USA, August 2009. AIAA 2009-5738.

Irene Gregory, Enric Xargay, Chengyu Cao, and Naira Hovakimyan. Flight test of an $\mathcal{L}_1$ adaptive controller on the NASA AirSTAR flight test vehicle. In *Proceedings of AIAA Guidance, Navigation, and Control Conference*, Toronto, Ontario, Canada, August 2010. AIAA 2010-8015.

Lukas Hewing, Juraj Kabzan, and Melanie N. Zeilinger. Cautious model predictive control using Gaussian process regression. *IEEE Transactions on Control Systems Technology*, 2019. Early Access.

Naira Hovakimyan and Chengyu Cao. $\mathcal{L}_1$ *Adaptive Control Theory: Guaranteed Robustness with Fast Adaptation*. SIAM, 2010.

Petros A. Ioannou and Jing Sun. *Robust Adaptive Control*. Dover Publications, Inc., Mineola, NY, 2012.

Hamidreza Jafarnejadsani, Donglei Sun, Hanmin Lee, and Naira Hovakimyan. Optimized $\mathcal{L}_1$ adaptive controller for trajectory tracking of an indoor quadrotor. *AIAA Journal of Guidance, Control, and Dynamics*, 40(6):1415–1427, 2017.

Rudolf E. Kalman. Design of a self-optimizing control system. *Transactions of the ASME*, 80: 468–478, 1958.

Isaac Kaminer, António Pascoal, Enric Xargay, Naira Hovakimyan, Chengyu Cao, and Vladimir Dobrokhodov. Path following for small unmanned aerial vehicles using $\mathcal{L}_1$ adaptive augmentation of commercial autopilots. *AIAA Journal of Guidance, Control, and Dynamics*, 33(2):550–564, 2010.

Isaac Kaminer, Enric Xargay, Venanzio Cichella, Naira Hovakimyan, António Manuel Pascoal, A. Pedro Aguiar, Vladimir Dobrokhodov, and Reza Ghabcheloo. Time-critical cooperative path following of multiple UAVs: Case studies. In *Advances in Estimation, Navigation, and Spacecraft Control*, pages 209–233. Springer, 2015.

Evgeny Kharisov, Kwang Ki Kevin Kim, Xiaofeng Wang, and Naira Hovakimyan. Limiting behavior of $\mathcal{L}_1$ adaptive controllers. In *Proceedings of AIAA Guidance, Navigation and Control Conference*, pages 451–456, Portland, Oregon, USA, August 2011. AIAA 2011-6441.

Yoan D. Landau. *Adaptive Control: The Model Reference Approach*. Control & Systems Theory. Marcel Dekker, Inc., New York, NY, 1979.

Armin Lederer, Jonas Umlauft, and Sandra Hirche. Uniform error bounds for Gaussian process regression with application to safe control. In *Proceedings of 33rd Conference on Neural Information Processing Systems*, pages 657–667, Vancouver, Canada, 2019.

Sergey Levine, Chelsea Finn, Trevor Darrell, and Pieter Abbeel. End-to-end training of deep visuomotor policies. *The Journal of Machine Learning Research*, 17(1):1334–1373, 2016.

Frank L. Lewis, Suresh Jagannathan, and Aydin Yeşildirek. *Neural Network Control of Robot Manipulators and Nonlinear Systems*. Taylor & Francis, Philadelphia, PA, 1998.

Timothy P. Lillicrap, Jonathan J. Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

Kumpati S. Narendra, Yuan-Hao Lin, and Lena S. Valavani. Stable adaptive controller design, part II: Proof of stability. *IEEE Transactions on Automatic Control*, 25(3):440–448, June 1980.

Yunpeng Pan, Ching-An Cheng, Kamil Saigol, Keuntaek Lee, Xinyan Yan, Evangelos A Theodorou, and Byron Boots. Imitation learning for agile autonomous driving. *The International Journal of Robotics Research*, 39(2-3):286–302, 2020.

Iman Salehi, Gang Yao, and Ashwin P. Dani. Active sampling based safe identification of dynamical systems using extreme learning machines and barrier certificates. In *Proceedings of 2019 International Conference on Robotics and Automation (ICRA)*, pages 22–28, Montreal, Canada, 2019.

Shankar Sastry and Marc Bodson. *Adaptive Control: Stability, Convergence and Robustness*. Dover Publications, Inc., Mineola, NY, 2011.

Ronald W. Schafer. What is a Savitzky-Golay filter? *IEEE Signal Processing Magazine*, 28(4): 111–117, 2011.

Steven Snyder. *$\mathcal{L}_1$ Adaptive Control within Learn-to-Fly*. PhD thesis, University of Illinois at Urbana-Champaign, 2019.

Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias W. Seeger. Information-theoretic regret bounds for Gaussian process optimization in the bandit setting. *IEEE Transactions on Information Theory*, 58(5):3250–3265, 2012.

Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT press, Cambridge, MA, 2018.

Andrew J. Taylor and Aaron D. Ames. Adaptive safety with control barrier functions. *arXiv preprint arXiv:1910.00555*, 2019.

Li Wang, Evangelos A. Theodorou, and Magnus Egerstedt. Safe learning of quadrotor dynamics using barrier certificates. In *Proceedings of 2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2460–2465, Brisbane, Australia, 2018.

Christopher K. I. Williams and Carl Edward Rasmussen. *Gaussian Processes for Machine Learning*. MIT press, Cambridge, MA, 2006.

Zongyu Zuo and Pengkai Ru. Augmented $\mathcal{L}_1$ adaptive tracking control of quadrotor unmanned aircrafts. *IEEE Transactions on Aerospace and Electronic Systems*, 50(4):3090–3101, 2014.