Differentially Flat Learning-Based Model Predictive Control Using a Stability, State, and Input Constraining Safety Filter

Adam W. Hall[1][†], Melissa Greeff[2][†], and Angela P. Schoellig[3][†]

Abstract—**Learning-based optimal control algorithms control unknown systems using past trajectory data and a learned model of the system dynamics. These controllers use either a linear approximation of the learned dynamics, trading performance for faster computation, or nonlinear optimization methods, which typically perform better but can limit real-time applicability. In this work, we present a novel nonlinear controller that exploits differential flatness to achieve similar performance to state-of-the-art learning-based controllers but with significantly less computational effort. Differential flatness is a property of dynamical systems whereby nonlinear systems can be exactly linearized through a nonlinear input mapping. Here, the nonlinear transformation is learned as a Gaussian process and is used in a safety filter that guarantees, with high probability, stability as well as input and flat state constraint satisfaction. This safety filter is then used to refine inputs from a flat model predictive controller to perform constrained nonlinear learning-based optimal control through two successive convex optimizations. We compare our method to state-of-the-art learning-based control strategies and achieve similar performance, but with significantly better computational efficiency, while also respecting flat state and input constraints, and guaranteeing stability.**

Index Terms—**Machine learning, Predictive control for nonlinear systems, Robotics.**

I. INTRODUCTION

IN recent years, interest has grown in controlling safety-critical systems whose dynamics are partially unknown, like unmanned aerial vehicles, driverless cars, and mobile manipulators. Classically, guaranteeing safety and stability of these uncertain systems results in overly conservative behaviour, limiting their usage in real tasks. Using machine learning and past trajectory data inside classical control frameworks to learn the system dynamics has proven to be an effective, safe learning-based control technique, but often requires slow nonlinear optimizations and can suffer from poor computational stability and efficiency [1]. For example, Gaussian process model predictive control (GPMPC) uses a Gaussian process (GP) to model the uncertain dynamics. This
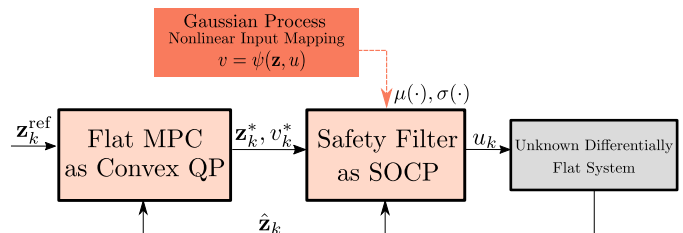

Fig. 1. Our proposed architecture enables high-performance trajectory tracking for uncertain differentially flat systems by developing a learning-based model predictive controller that is computationally efficient to compute. We do this by solving two convex optimization problems: 1) a convex quadratic program in flat model predictive control that finds the flat input v_k^* and state z_k^* to track a reference z_k^{ref}; 2) a second-order cone program in the safety filter that uses a learned representation of the flat nonlinear input mapping $v = \psi(z, u)$, with posterior prediction mean $\mu(\cdot)$ and covariance $\sigma(\cdot)$, to perform probabilistic feedback linearization guaranteeing asymptotic stability, as well as state and input constraint satisfaction.

learned model is then used inside a robust model predictive control (MPC) framework. GPMPC has been used on mobile robots [2], quadrotors [3], and other autonomous systems.

A drawback of GPMPC is its computational complexity. It either requires powerful on-board computation or remote computation of inputs, limiting its use in real systems. Standard GPs require all of their training data to be stored in memory and their posterior mean and covariance predictions are computationally expensive, even when using approximate methods [4]—many approximations and 'tricks' are used to achieve real-time operability. There are, however, structural assumptions about the dynamics that can improve computational speed without sacrificing performance, such as incorporating a control-affine structure, and differential flatness.

Differential flatness is a property of many nonlinear dynamical systems that enables their transformation into linear systems through a nonlinear input mapping, called exact linearization [5]—this is an exact transformation, not an approximation. Linear control techniques can then be used to compute a flat input that is subsequently transformed through this nonlinear mapping and applied to the real system. Many real robotic systems are differentially flat, like quadrotors [6], flexible-joint manipulators [7], and mobile robots [8], to list a few. In the literature, differential flatness has been exploited to control nonlinear systems using linear MPC [6], called flat MPC (FMPC), which achieves similar performance to nonlinear MPC while greatly improving computational efficiency.

Differential flatness has also been exploited in safe learning-based controllers. In [9], the nonlinear input mapping and its inverse are learned from trajectory data and used in a robust linear quadratic regulator (RLQR) formulation to control an uncertain system, while guaranteeing an ultimate upper

[1]Adam W. Hall is jointly with the Learning Systems and Robotics Lab (www.learnsyslab.org) and the STARS lab (starslab.ca) at the University of Toronto Institute for Aerospace Studies (UTIAS), Toronto, Canada. Email: adam.hall@robotics.utias.utoronto.ca

[2]Melissa Greeff is with the Robora Lab (www.roboralab.com) at the Department of Electrical and Computer Engineering, Queen's University, Kingston, Canada.

[3]Angela P. Schoellig is with the Learning Systems and Robotics Lab (www.learnsyslab.org) at the Technical University of Munich and the University of Toronto, and the Munich Institute for Robotics and Machine Intelligence (MIRMI), Munich, Germany.

[†]All authors are with the Vector Institute for Artificial Intelligence, Toronto, Canada.

bound on tracking error. This method is more computationally efficient than GPMPC, however, still involves a demanding nonlinear optimization and cannot explicitly handle input constraints or state constraints.

More recently, [10] has used a learned representation of the nonlinear input mapping inside of a safety filter that finds the input that best matches the desired input from a flat linear controller while guaranteeing probabilistic asymptotic stability. In particular, the safety filter is formulated as a Second-Order Cone Program (SOCP) by exploiting differential flatness and the affine-control structure of the dynamics. This SOCP is more computationally efficient than GPMPC and RLQR and can explicitly handle input constraints, but it cannot enforce any state constraint guarantees.

In this work, we build upon [10] to provide three main contributions:

- a novel safe learning-based MPC for nonlinear, differentially flat, control-affine systems that can be solved via two successive convex optimizations;
- a novel safety filter that guarantees, with high probability, asymptotically stable tracking error, and flat-state and input constraint satisfaction, modelled as an SOCP;
- a comparison with GPMPC, in simulation, that shows our approach achieves similar performance but is at least 10 times more computationally efficient.

The advantage of MPC, used in this work, over linear quadratic regulation (LQR) used in [10], is anticipating the reference and constraint boundaries which avoids infeasible states and aggressive inputs, as shown in our simulated examples. All code is available at github.com/utiasDSL/fmpc_socp.

II. PROBLEM STATEMENT

We consider a single-input continuous-time control-affine nonlinear system

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t)) + g(\mathbf{x}(t))u(t), \tag{1}$$

with initial condition $\mathbf{x}(0) = \mathbf{x}_0$, where $\mathbf{x}(t) \in \mathbb{R}^n$ and $u(t) \in \mathbb{R}$ are the state and input of the system at time $t \in \mathbb{R}_{\geq 0}$. The maps $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^n$ are unknown, but are assumed to be locally Lipschitz continuous. It is assumed that a prior model of the unknown system is given by

$$\dot{\mathbf{x}}(t) = \tilde{f}(\mathbf{x}(t)) + \tilde{g}(\mathbf{x}(t))u(t). \tag{2}$$

Assumption 2.1: The system (1) and the prior model (2) are single-input single-output systems and differentially flat with respect to a *known* flat output $y = h(\mathbf{x}(t))$, with $y \in \mathbb{R}$.

Definition 2.1 (Differential Flatness [5]): A smooth single input nonlinear system is *differentially flat* if there exists a *flat output* $y \in \mathbb{R}$ such that the input and all states can be uniquely determined from this *flat output* and its derivatives $\mathbf{z} = [y, \dot{y}, \ldots, y^{(n-1)}]^T$, and there exists smooth, invertible functions $\mathbf{x} = \phi^{-1}(\mathbf{z})$, $u = \psi^{-1}(\mathbf{z}, v)$, where $v = y^{(n)}$.

Lemma 2.1 (Linearized Flat Dynamics [5], [7]): A differentially flat system (1) can be transformed into a linear system in the Brunowsky Canonical form

$$\dot{\mathbf{z}}(t) = \mathbf{A}\,\mathbf{z}(t) + \mathbf{B}\,v(t), \tag{3}$$

where the new input v is related to the system input via the nonlinear mapping

$$v = \psi(\mathbf{z}, u). \tag{4}$$

Furthermore, if the system is control affine, then this mapping has the specific form

$$v = \alpha(\mathbf{z}) + \beta(\mathbf{z})u. \tag{5}$$

Remark 2.1: Given that the mappings f and g in (1) are unknown, α and β are also unknown.

Our system (1) is subject to input constraints $\mathcal{U} := \{u \in \mathbb{R} \mid u_{\min} \leq u \leq u_{\max}\}$ and convex constraints on the flat-state vector $\mathbf{z} \in \mathcal{Z}$.

Remark 2.2: The constraints are enforced on the flat state $\mathbf{z}$. While the flat state differs from the state $\mathbf{x}$, in many robotic systems it still represents physical quantities that are constrained. For example, in quadrotors, it represents the position, velocity, and acceleration; in flexible joint manipulators, it represents the output shaft joint angle and its derivatives [7]; and for mobile robotics, the commonly used unicycle model's flat state comprises position, velocity, and acceleration [8].

The objective is to design a computationally efficient controller for the unknown system (1) that achieves high tracking performance of a reference trajectory $\mathbf{z}^{\mathrm{ref}}$, guarantees tracking convergence with high probability, and respects input and state constraints.

III. BACKGROUND

A. Discrete-Time Control Lyapunov Function

Given a constant sampling period δ_t, the discretization of (3) becomes

$$\mathbf{z}_{k+1} = \mathbf{A}_{\mathrm{d}}\,\mathbf{z}_k + \mathbf{B}_{\mathrm{d}}\,v_k, \tag{6}$$

where $\mathbf{A}_{\mathrm{d}}$ and $\mathbf{B}_{\mathrm{d}}$ are Euler discretizations of $\mathbf{A}$ and $\mathbf{B}$. Further, $\mathbf{z}_k = z(\delta_t k)$ and $v_k = v(\delta_t k)$ are time-sampled flat states and inputs at time step $k \in \mathbb{N}_{\geq 0}$.

Consider the smooth reference $\mathbf{z}^{\mathrm{ref}}(t) : [0, T] \to \mathbb{R}^n$ and $v^{\mathrm{ref}}(t) : [0, T] \to \mathbb{R}$ sampled every $k\delta_t$ to yield the discrete reference signals $\mathbf{z}_k^{\mathrm{ref}} = \mathbf{z}^{\mathrm{ref}}(k\delta_t)$, $v_k^{\mathrm{ref}} = v^{\mathrm{ref}}(k\delta_t)$. The tracking error can be defined with respect to this reference as $\mathbf{e}_k = \mathbf{z}_k - \mathbf{z}_k^{\mathrm{ref}}$. Given an error feedback control policy $v_k = -\mathbf{K}\,\mathbf{e}_k + v_k^{\mathrm{ref}}$ with gain $\mathbf{K} \in \mathbb{R}^{1 \times n}$, the error dynamics become $\mathbf{e}_{k+1} = (\mathbf{A}_{\mathrm{d}} - \mathbf{B}_{\mathrm{d}}\,\mathbf{K})\mathbf{e}_k$.

Definition 3.1: If a function $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ satisfies

$$V(0) = 0 \text{ and } V(\mathbf{e}_k) > 0,\ \forall \mathbf{e}_k \in \mathcal{Z} \setminus \{0\}$$
$$V(\mathbf{e}_{k+1}) < V(\mathbf{e}_k),\ \forall \mathbf{e}_k \in \mathcal{Z} \setminus \{0\} \tag{7}$$

for the error dynamics, then it is called a *Control Lyapunov Function* (CLF) and its existence guarantees the asymptotic stability of the closed-loop dynamics.

Lemma 3.1 ([7]): Given Assumption 2.1, if the transformation (4) is known, u_k can be chosen to cancel the nonlinear term (4) and find a v_k such that the resulting linear error dynamics are Hurwitz.

Remark 3.1: The true dynamics are unknown, and thus the inverse transform $u = \psi(\mathbf{z}, v)$ is unknown and is often approximated using the prior model (2). Such control methods, however, are not robust to model mismatch in (4) [6].

B. Gaussian Processes (GPs)

GPs are used to model nonlinear functions $\psi(\mathbf{a})$: $\mathbb{R}^{\dim(\mathbf{a})} \to \mathbb{R}$. They encapsulate a prior over possible functions. As data are collected, the possible functions for $\psi(\mathbf{a})$ are refined and the GP obtains a posterior distribution over functions [11]. GPs assume that all collected data is jointly Gaussian with a prior mean and covariance. GP regression is the process of finding the hyperparameters that optimize the log-likelihood of the marginal distribution over the sampled function data.

Given a query point $\mathbf{a}^*$, the posterior prediction conditioned on the data $\mathcal{D} = \{(\mathbf{a}_i, \hat{\psi}(\mathbf{a}_i)\}_{i=1}^{N_{\mathcal{D}}}$ is given by the distribution $\psi(\mathbf{a}^*)|\mathcal{D} \sim \mathcal{N}\left(\mu(\mathbf{a}^*), \sigma^2(\mathbf{a}^*)\right)$. Here, the posterior mean and covariance are $\mu(\mathbf{a}^*) = \mathbf{k}(\mathbf{a}^*)\mathbf{K}^{-1}\hat{\Psi}$, and $\sigma(\mathbf{a}^*) = k(\mathbf{a}^*, \mathbf{a}^*) - \mathbf{k}(\mathbf{a}^*)\mathbf{K}^{-1}\mathbf{k}^T(\mathbf{a}^*)$, where $\mathbf{k}(\mathbf{a}^*) = [k(\mathbf{a}^*, \mathbf{a}_i), \ldots, k(\mathbf{a}^*, \mathbf{a}_{N_{\mathcal{D}}})]$, $\mathbf{K}_{i,j} = k(\mathbf{a}_i, \mathbf{a}_j)$, and $\hat{\Psi} = [\hat{\psi}(\mathbf{a}_1), \ldots, \hat{\psi}(\mathbf{a}_{N_{\mathcal{D}}})]^T$. See [11] for more details.

IV. METHODOLOGY

In our proposed architecture, shown in Figure 1, we exploit the control-affine and differentially flat structure of the system (1) to design a learning-based model predictive controller. Our method solves two convex optimization problems at each time-step: a convex FMPC to determine the flat input to the linear dynamics (6) that best tracks a given reference trajectory; and a safety filter formulated as an SOCP that ensures probabilistic asymptotic stability, probabilistic flat-state constraint satisfaction, and input constraint satisfaction, even when (1) is uncertain. By leveraging system data, we develop a controller that achieves high tracking performance despite unknown system dynamics (1), but is still computationally efficient, making onboard input computation more practical relative to the current state-of-the-art.

This section follows the components of our controller given in Figure 1. First, in Section IV-A, the FMPC is formulated, assuming the system dynamics (1) and the input mapping $u = \psi^{-1}(\mathbf{z}, v)$ are known. As (1) is not known, a learned representation of (4) using GPs is used. In particular, our approach exploits the affine structure of (1) and (4), as detailed in Section IV-B. Finally, in Section IV-C, the probabilistic feedback linearization (Section IV-C.1), probabilistic asymptotic stability (Section IV-C.2), and probabilistic flat-state constraint (Section IV-C.3), are formulated as an SOCP safety filter (Section IV-C.4).

A. Flat Model Predictive Control

FMPC iteratively solves a convex finite-horizon optimal control problem (OCP) to control the nonlinear flat system (1). In this section, the FMPC formulation is presented and builds on the FMPC formulation in [6]. We highlight how FMPC can be designed such that the error dynamics, with respect to a flat reference $\mathbf{z}^{\text{ref}}$, are asymptotically stable when the system dynamics (1) are known.

Assumption 4.1 ([12]): There exists a cost function ℓ that is bounded by comparison functions $\bar{\zeta}, \underline{\zeta} \in \mathcal{K}_\infty$ such that $\underline{\zeta}(\|\mathbf{z}_k - \mathbf{z}_k^{\text{ref}}\|) \leq \inf_{v_k} \ell(\mathbf{z}_k, \mathbf{z}_k^{\text{ref}}, v_k) \leq \bar{\zeta}(\|\mathbf{z}_k - \mathbf{z}_k^{\text{ref}}\|) \ \forall \ \mathbf{z}_k \in \bar{\mathcal{Z}}$.

Remark 4.1: We consider cost functions of the form $\ell(\mathbf{z}_k, \mathbf{z}_k^{\text{ref}}, v_k) = (\mathbf{z}_k - \mathbf{z}_k^{\text{ref}})^T \mathbf{Q}(\mathbf{z}_k - \mathbf{z}_k^{\text{ref}}) + v_k^2 r$, where $\mathbf{Q} \in \mathbb{R}^{n \times n}$, $\mathbf{Q} \succ 0$, $r \in \mathbb{R}$, $r > 0$.

The convex finite-horizon OCP solved at each time step is given by

$$\min_{\mathbf{z}_{k|k:k+N}, v_{k|k:k+N-1}} J_N(\mathbf{z}_{k|k+1:k+N}, \mathbf{z}_{k|k+1:k+N}^{\text{ref}}, v_{k|k:k+N-1})$$

$$\text{s.t.} \quad \mathbf{z}_{k|k} = \hat{\mathbf{z}}_k \tag{8}$$

$$\mathbf{z}_{k|i+1} = \mathbf{A}_{\text{d}}\,\mathbf{z}_{k|i} + \mathbf{B}_{\text{d}}\,v_{k|i}, \ \forall i \in [k, k+N-1]$$

$$\mathbf{z}_{k|i} \in \mathcal{Z} \ \forall i \in [k, k+N],$$

where $N \in \mathbb{N}$ is the horizon length, $\hat{\mathbf{z}}_k$ is the measured flat state at time step k, $J_N(\cdot, \cdot, \cdot) = \sum_{i=k+1}^{k+N} \ell(\mathbf{z}_{k|i}, \mathbf{z}_{k|i}^{\text{ref}}, v_{k|i-1})$ is the cost function, and the flat states $\mathbf{z}_{k|k:k+N} = [\mathbf{z}_k, \ldots, \mathbf{z}_{k+N}]$, the flat reference $\mathbf{z}_{k|k:k+N}^{\text{ref}} = [\mathbf{z}_k^{\text{ref}}, \ldots, \mathbf{z}_{k+N}^{\text{ref}}]$, and the flat inputs $v_{k|k:k+N-1} = [v_k, \ldots, v_{k+N-1}]$ are sequences for time step k to time step $k+N$, computed at time step k. The optimal solution to OCP (8) at time step k is given by the sequences $\mathbf{z}_{k|k:k+N}^*$ and $v_{k|k:k+N-1}^*$. We use the notation $\mathbf{z}_k^* := \mathbf{z}_{k|k}^*$ and $v_k^* := v_{k|k}^*$ for the optimal flat state and flat input computed at time step k. This state-input pair is used in the safety filter in Section IV-C. This OCP then applies the zero-order hold input $u(t) = \psi^{-1}(\mathbf{z}_k^*, v_k^*) \ \forall \ t \in [k\delta_t, (k+1)\delta_t)$ to the system.

Assumption 4.2 ([12]): There exists a comparison function $\xi \in \mathcal{K}_\infty$ and an integer $s \in \mathbb{N}$ such that, for all $\mathbf{z} \in \mathcal{Z}$, the inequality $\inf_v J_s(\mathbf{z}, \mathbf{z}^{\text{ref}}, v) \leq \xi(\inf_v \ell(\mathbf{z}, \mathbf{z}^{\text{ref}}, v))$ holds for all $s \in \mathbb{N}$.

Lemma 4.1 ([12, Thm. 6.2], [13, Sec. 8.3]): Given Assumptions 4.1 and 4.2, there exists an $N \in \mathbb{N}$ such that the sampled-data MPC, defined by OCP (8), is *asymptotically stable* in the closed-loop with respect to the reference $\mathbf{z}^{\text{ref}}$. Furthermore, in the unconstrained case (i.e., where the flat-state constraint $\mathbf{z}_i \in \mathcal{Z}$ is not present), the OCP (8) permits an equivalent closed-form solution

$$v_k^* = -\mathbf{K}(\mathbf{z}_k - \mathbf{z}_k^{\text{ref}}) + v_k^{\text{ref}}, \tag{9}$$

where $\mathbf{K} \in \mathbb{R}^{1 \times n}$ is the equivalent gain matrix.

B. Gaussian Process Learning

FMPC, presented in the previous section, relies on knowing (1) to compute the input $u_k = \psi^{-1}(\mathbf{z}_k^*, v_k^*)$ from the optimized trajectory (i.e., $\mathbf{z}_k^*$ and v_k^*). Given (1) is unknown, we propose to learn the map (4) as a GP. We thus encode the control-affine structure of (5) in the kernel selection of the GP. This structure enables us to formulate probabilistic stability and state constraints in our filter in Section IV-C as an SOCP. Specifically, we select the following kernel [14]

$$k(\mathbf{a}_i, \mathbf{a}_j) = k_\alpha(\mathbf{z}_i, \mathbf{z}_j) + u_i k_\beta(\mathbf{z}_i, \mathbf{z}_j)u_j + \delta_{i,j}\sigma_\eta^2. \tag{10}$$

Assumption 4.3: $k_\alpha(\cdot, \cdot)$ and $k_\beta(\cdot, \cdot)$ are positive definite and bounded kernels.

Lemma 4.2: [10, Lem. 2] Given Assumption (4.3), the affine kernel in (10) is also bounded and positive definite. Given a query point $\mathbf{a} = (\mathbf{z}, u)$ and a regressed GP conditioned on $N_{\mathcal{D}}$ noisy observations $\mathcal{D} = \{\mathbf{a}_i, \hat{\psi}(\mathbf{a}_i)\}_{i=1}^{N_{\mathcal{D}}}$, the

posterior mean prediction $\mu(\mathbf{a})$ and variance $\sigma(\mathbf{a})$ are linear and quadratic in u, respectively,

$$\mu(\mathbf{a}) = \gamma_1(\mathbf{z}) + \gamma_2(\mathbf{z})u, \tag{11}$$

$$\sigma^2(\mathbf{a}) = \gamma_3(\mathbf{z}) + \gamma_4(\mathbf{z})u + \gamma_5(\mathbf{z})u^2, \tag{12}$$

where

$$\gamma_1(\mathbf{z}) = \mathbf{k}_\alpha(\mathbf{z})\mathbf{K}^{-1}\hat{\mathbf{\Psi}}, \qquad \gamma_2(\mathbf{z}) = \mathbf{k}_\beta(\mathbf{z})\mathbf{K}^{-1}\hat{\mathbf{\Psi}},$$
$$\gamma_3(\mathbf{z}) = k_\alpha(\mathbf{z}, \mathbf{z}) - \mathbf{k}_\alpha(\mathbf{z})\mathbf{K}^{-1}\mathbf{k}_\alpha^T(\mathbf{z}),$$
$$\gamma_4(\mathbf{z}) = -(\mathbf{k}_\beta(\mathbf{z})\mathbf{K}^{-1}\mathbf{k}_\alpha^T(\mathbf{z}) + \mathbf{k}_\alpha(\mathbf{z})\mathbf{K}^{-1}\mathbf{k}_\beta^T(\mathbf{z})),$$
$$\gamma_5(\mathbf{z}) = k_\beta(\mathbf{z}, \mathbf{z}) - \mathbf{k}_\beta(\mathbf{z})\mathbf{K}^{-1}\mathbf{k}_\beta^T(\mathbf{z}).$$

Here, $\hat{\mathbf{\Psi}} \in \mathbb{R}^{N_\mathcal{D}}$, with $\hat{\mathbf{\Psi}}_i = \hat{\psi}(\mathbf{a}_i)$, $\mathbf{k}_\alpha(\mathbf{z}) \in \mathbb{R}^{1 \times N_\mathcal{D}}$ with $\mathbf{k}_{\alpha,i}(\mathbf{z}) = k_\alpha(\mathbf{z}, \mathbf{z}_i)$, $\mathbf{k}_\beta(\mathbf{z}) \in \mathbb{R}^{1 \times N_\mathcal{D}}$ with $\mathbf{k}_{\beta,i}(\mathbf{z}) = k_\beta(\mathbf{z}, \mathbf{z}_i)$, and $\mathbf{K} \in \mathbb{R}^{N_\mathcal{D} \times N_\mathcal{D}}$ with elements $\mathbf{K}_{i,j} = k(\mathbf{a}_i, \mathbf{a}_j)$.

C. Safety Filter

In this section, we use the learned GP model of (5) in a safety filter design with three components: 1) probabilistic feedback linearization, 2) a probabilistic stability constraint, and 3) a probabilistic state constraint. We then show how to implement the safety filter as a SOCP.

1) Probabilistic Feedback Linearization: We aim to find u_k such that the flat input v_k in (6) seen by the system closely matches the desired flat input v_k^*, optimized in FMPC. We select the input u_k that minimizes the expected square distance between the desired flat input v_k^* and the output of the GP model for (4) as

$$\min_{u_k} \mathbb{E}\left[\|\psi(\mathbf{z}_k^*, u_k) - v_k^*\|^2\right]. \tag{13}$$

When we query the GP model of (1) at $\mathbf{a} = (\mathbf{z}_k^*, u_k)$ the posterior prediction of (4) is normally distributed $\psi(\mathbf{z}_k^*, u_k)|\mathcal{D} = \mathcal{N}\left(\mu(\mathbf{z}_k^*, u_k), \sigma^2(\mathbf{z}_k^*, u_k)\right)$. Consequently, (13) can be written as $\min_{u_k}(\mu(\mathbf{z}_k^*, u_k) - v_k^*)^2 + \sigma^2(\mathbf{z}_k^*, u_k)$. Exploiting the affine form of the GP kernel selection allows for the mean and covariance to be substituted by (11) and (12), respectively, further simplifying (13) to

$$\min_{u_k}(\gamma_2^{*2} + \gamma_5^*)u_k^2 + (2\gamma_1^*\gamma_2^* - 2\gamma_2^*v^* + \gamma_4^*)u_k, \tag{14}$$

where $\gamma_i^* := \gamma_i(\mathbf{z}_k^*)$.

Remark 4.2: Following from [10], the optimization problem in (14) is convex since $\gamma_5^* \geq 0$ as it is the predicted covariance of $\beta(\mathbf{z})$ in (5), and the minimization is quadratic in the optimization variable u_k.

2) Probabilistic Stability Constraints: We formulate a probabilistic stability constraint using the CLF from (7) to ensure that the input u_k guarantees probabilistic stability for the closed-loop system, despite (1) being unknown.

To formulate this constraint, consider the Lyapunov function of the form $V(\mathbf{e}_k) = \mathbf{e}_k^T \mathbf{P} \mathbf{e}_k$ and a nominal flat input $v_k^{\text{nom}} = -\mathbf{K}\mathbf{e}_k + v_k^{\text{ref}}$, using the gain computed in (9). Then, the error at $k+1$ can be expressed as $\mathbf{e}_{k+1} = \mathbf{A}_\text{d}\mathbf{e}_k - \mathbf{B}_\text{d}\mathbf{K}\mathbf{e}_k + \mathbf{B}_\text{d}(\psi(\mathbf{z}_k^*, u_k) - v_k^{\text{nom}})$. After using the discrete-time algebraic Ricatti equation, the CLF decrease condition

(7) can be expressed as

$$\mathbf{e}_k^T[\mathbf{P} - \mathbf{Q} - r\,\mathbf{K}^T\mathbf{K}]\mathbf{e}_k$$
$$-2\mathbf{e}_k^T(\mathbf{A}_\text{d} - \mathbf{B}_\text{d}\mathbf{K})^T\mathbf{P}\mathbf{B}_\text{d}(\psi(\mathbf{z}_k^*, u_k) - v_k^{\text{nom}}) \tag{15}$$
$$+(\psi(\mathbf{z}_k^*, u_k) - v_k^{\text{nom}})^2\mathbf{B}_\text{d}^T\mathbf{P}\mathbf{B}_\text{d} \leq \mathbf{e}_k^T\mathbf{P}\mathbf{e}_k - \epsilon,$$

where $\epsilon > 0$ is a small constant to allow for the inequality to be non-strict. Thus, the left-hand side of (15) is quadratic in $\psi(\mathbf{z}_k^*, u_k)$ which we have learned as a GP. We conservatively bound the last term $(\psi(\mathbf{z}_k^*, u_k) - v_k^{\text{nom}})^2\mathbf{B}_\text{d}^T\mathbf{P}\mathbf{B}_\text{d}$ as the eigenvalues of $\mathbf{B}_\text{d}^T\mathbf{P}\mathbf{B}_\text{d}$ are proportional to δ_t^2, making this term small relative to the other terms. Using the fact that the posterior mean prediction (11) is affine in u_k and that u_k is bounded $u_{\min} \leq u_k \leq u_{\max}$, (15) can be written as

$$-w_1(\psi(\mathbf{z}_k^*, u_k) - v_k^{\text{nom}}) \leq w_3 - w_2, \tag{16}$$

where

$$w_1 := 2\mathbf{e}_k^T(\mathbf{A}_\text{d} - \mathbf{B}_\text{d}\mathbf{K})^T\mathbf{P}\mathbf{B}_\text{d},$$
$$w_2 := \mathbf{B}_\text{d}^T\mathbf{P}\mathbf{B}_\text{d} \max_{s=\{u_{\min}, u_{\max}\}}\|\mu(\mathbf{z}_k^*, s) - v_k^{\text{nom}}\|^2,$$
$$w_3 := \mathbf{e}_k^T[\mathbf{Q} + r\,\mathbf{K}^T\mathbf{K}]\mathbf{e}_k - \epsilon.$$

Assumption 4.4: The nonlinear single-input control affine system (1) permits a bounded reproducible kernel Hilbert space (RKHS) norm $\|\psi(\mathbf{z}_k, u_k)\|_{\text{kern}}$ with respect to the kernel (10) used in the GP, and the GP's observation noise η is uniformly bounded by σ_η.

Lemma 4.3: Let $\delta \in (0, 1)$. Given Assumption 4.4, $\Pr\{-w_1(\mu(\mathbf{z}_k^*, u_k) - v_k^{\text{nom}}) \leq w_3 - w_2 - |w_1|\beta^{1/2}\sigma(\mathbf{z}_k^*, u_k)\} \geq 1 - \delta$, where $\beta = 2\|\psi(\mathbf{z}_k^*, u_k)\|_{\text{kern}} + 300\gamma\ln^3((N+1)/\delta)$, and $\gamma \in \mathbb{R}$ is the maximum information gain.

Proof: Given Assumption 4.4 and Theorem 3 in [15], the mean prediction of a GP is bounded with respect to the true function evaluation as per $\Pr\{\forall\mathbf{a} \in \mathcal{A}, |\psi(\mathbf{a}) - \mu(\mathbf{a})| \leq \beta^{1/2}\sigma(\mathbf{a})\} \geq 1 - \delta$. When considering this probabilistic bound in the context of the left-hand side of (16) and expanding the absolute value, the upper inequality bound becomes $-w_1(\psi(\mathbf{z}_k^*, u_k) - v_k^{\text{nom}}) \leq -w_1(\mu(\mathbf{z}_k^*, u_k) - v_k^{\text{nom}}) - |w_1|\beta^{1/2}\sigma(\mathbf{z}_k^*, u_k)$. Using this expression with Theorem 3 from [15] means that the inequality holds true with probability $1 - \delta$. Using the probabilistic bound in (16), yields the probabilistic constraint given in the lemma. ∎

3) Probabilistic State Constraints: We also ensure that $\mathbf{z}_{k+1} \in \mathcal{Z}$, with high probability, by taking into account the uncertainty in the prediction of $\mathbf{z}_{k+1}$ due to the uncertainty in the learned mapping $\psi(\mathbf{z}_k^*, u_k)$. A constraint is required here beyond the state constraint in (8) as the input can be modified by the safety filter, and the uncertainty in $\psi(\mathbf{z}_k^*, u_k)$ must be accounted for. We thus first determine the uncertainty in $\mathbf{z}_{k+1}$, then use it to tighten $\mathcal{Z}$ to ensure $\mathbf{z}_{k+1} \in \mathcal{Z}$.

Using the posterior mean prediction from the GP in (6), the mean of the next state becomes $\mu_{k+1}^z = \mathbf{A}_\text{d}\mathbf{z}_k^* + \mathbf{B}_\text{d}\mu(\mathbf{z}_k^*, u_k)$. For brevity, we only consider uncertainty in μ_{k+1}^z due to $\sigma(\mathbf{z}_k^*, u_k)$. Thus, the uncertainty in μ_{k+1}^z is given as $\sigma_{\mathbf{z}_{k+1}}^2 = \mathbf{B}_\text{d}\sigma^2(\mathbf{z}_k^*, u_k)\mathbf{B}_\text{d}^\top$, which we use to tighten the constraint set $\mathcal{Z}$ using probabilistic reachable sets (PRS).

Definition 4.1 (One-Step PRS [16]): Given the residual error of a random sample away from its mean $\Delta_k = \mu_k^Z - \mathbf{z}_k$,

a set $\mathcal{R}$ is called a One-Step PRS of probability level δ if $\Pr(\Delta_{k+1} \in \mathcal{R} | \Delta_k = 0) \geq \delta$.

Remark 4.3: If we define an error term $\Delta_{k+1}^z = \mu_{k+1}^z - \mathbf{z}_{k+1}$, we can define the tightened constraints on μ_{k+1}^z as

$$\mu_{k+1}^z \in \mathcal{Z} \ominus \mathcal{R}(\sigma_{\mathbf{z}_{k+1}}), \tag{17}$$

where $\ominus$ represents the Pontryagin set difference.

Lemma 4.4 (Probabilistic Half-Space Constraints): Consider a half-space constraint given by $\mathcal{Z}^{hs} := \{\mathbf{z}_k | \mathbf{h}^T \mathbf{z}_k \leq b\}$ with $\mathbf{h} \in \mathbb{R}^n$ and $b \in \mathbb{R}_+$ defining the constraint. Then, given the uncertainty in the dynamics propagation, the tightened constraint becomes

$$\mathcal{Z}^{hs}(\sigma_{\mathbf{z}_{k+1}}) := \left\{ \mathbf{z}_{k+1} | \mathbf{h}^T \mathbf{z}_{k+1} \leq b - \rho(\delta) \sqrt{\mathbf{h}^T \sigma_{\mathbf{z}_{k+1}}^2 \mathbf{h}} \right\},$$

which guarantees that the constraint will be satisfied given the uncertainty in the dynamics, with probability level δ. Here, ρ is the quantile function of a standard Gaussian random variable.

Proof: Given the uncertainty in the next state, $\sigma_{\mathbf{z}_{k+1}}$ and under the random variable Δ_{k+1}^z, the marginal distribution becomes $\mathbf{h}^T \Delta_{k+1}^z \sim \mathcal{N}\left(0, \mathbf{h}^T \sigma_{\mathbf{z}_{k+1}}^2 \mathbf{h}\right)$. Using the quantile function of a standard Gaussian $\rho(\delta)$ with probability level δ a PRS can be constructed

$$\mathcal{R}^z(\sigma_z) := \left\{ \Delta | \mathbf{h}^T \Delta_{k+1}^z \leq \rho(\delta) \sqrt{\mathbf{h}^T \sigma_{\mathbf{z}_{k+1}}^2 \mathbf{h}} \right\}, \tag{18}$$

as follows from [16]. Using (18) in (17), the probabilistic half-space constraint is as shown in Lemma 4.4. ∎

4) Safety Filter as an SOCP: Here, we formulate the safety filter as an SOCP.

Theorem 4.1 (Second-Order Cone Program): Given Assumptions 2.1, 4.1, 4.2, 4.3, and 4.4, the optimization problem given in (14) subject to the probabilistic asymptotic stability constraint given by Lemma 4.3 and probabilistic state constraint Lemma 4.4 can be written as an SOCP

$$\min_{\bar{\mathbf{u}}} \quad [2\gamma_1^* \gamma_2^* - 2\gamma_2^* v_k^* + \gamma_4^*, 1] \bar{\mathbf{u}},$$
$$\text{s.t.} \quad \|\bar{\mathbf{A}}_i \bar{\mathbf{u}} + \bar{\mathbf{b}}_i\| \leq \bar{\mathbf{c}}_i^T \bar{\mathbf{u}} + \bar{\mathbf{d}}_i \quad i \in \{1, 2, 3\}, \tag{19}$$
$$u_{\min} \leq u_k \leq u_{\max},$$

where $\bar{\mathbf{u}} = [u_k, q]^T$, $\bar{\mathbf{A}}_i \in \mathbb{R}^{2 \times 2}$, $\bar{\mathbf{b}}_i \in \mathbb{R}^2$, $\bar{\mathbf{c}}_i \in \mathbb{R}^2$, $\bar{\mathbf{d}}_i \in \mathbb{R}$.

Proof: First, a dummy variable $q \geq (\gamma_2^{*2} + \gamma_5^*) u_k^2$ is introduced into the optimization problem to reformulate the cost as linear in $\bar{\mathbf{u}}$. Moreover, $0 \geq (\gamma_2^{*2} + \gamma_5^*) u_k^2 - 4q$ which can be rewritten as $(1 + q)^2 \geq 4(\gamma_2^{*2} + \gamma_5^*) u^2 + (1 - q)^2$. Note that both sides of this inequality are positive, so this can be rewritten as a standard SOC constraint where $\bar{\mathbf{A}}_1 = \text{diag}\left\{2\sqrt{\gamma_2^{*2} + \gamma_5^*}, -1\right\}$, $\bar{\mathbf{b}}_1 = [0, 1]^T$, $\bar{\mathbf{c}}_1 = [0, 1]^T$, and $\bar{\mathbf{d}}_1 = 1$. The optimization objective then becomes $[2\gamma_1^* \gamma_2^* - 2\gamma_2^* v^* + \gamma_4^*, 1] \bar{\mathbf{u}}$.

Next, consider the probabilistic stability constraint (16). Using the posterior prediction mean and covariance expressions (11) and (12), it is noted that

$$\sqrt{\gamma_3^* + \gamma_4^* u_k + \gamma_5^* u_k^2} = \left\| \begin{bmatrix} \sqrt{\gamma_5^*} u_k + \frac{\gamma_4^*}{2\sqrt{\gamma_5^*}} \\ \sqrt{\gamma_3^* - \frac{\gamma_4^{*2}}{4\gamma_5^*}} \end{bmatrix} \right\|_2. \tag{20}$$

Using this, the stability constraint can be rewritten as an SOC constraint with $\bar{\mathbf{A}}_2 = \text{diag}\left\{w_1\sqrt{\gamma_5^*}, 0\right\}$, the vectors

$k \to 0$.
if *Training offline* **then**
| Train the nonlinear map $v = \psi(\mathbf{z}, u)$.
while $k\delta_t \leq T$ **do**
| Measure the current flat state $\hat{\mathbf{z}}_k$.
| Find the optimal flat state and input $\mathbf{z}_k^*, v_k^*$ by
| solving the OCP (8).
| Find u_k^* that minimizes (19) using $\mathbf{z}_k^*$ and v_k^*.
| Apply u_k^* to the real system.
| Set $k \leftarrow k + 1$.
| **if** *Training online* **then**
| | Update the nonlinear map with measured data.
end

Algorithm 1: Proposed control algorithm.

$\bar{\mathbf{b}}_2 = \left[w_1 \frac{\gamma_4^*}{2\sqrt{\gamma_5^*}}, w_1 \sqrt{\gamma_3^* - \frac{\gamma_4^{*2}}{4\gamma_5^*}} \right]^T$ and $\bar{\mathbf{c}}_2 = [\frac{w_1 \gamma_5^*}{\beta^{1/2}}, 0]^T$, and the scalar $\bar{\mathbf{d}}_2 = (w_1 \gamma_2^* + w_3 - w_2)/\beta^{1/2}$.

Finally, the probabilistic state constraint from Lemma 4.4 is transformed into an SOC constraint. First, see that the term $\sqrt{\mathbf{h}^T \sigma_{\mathbf{z}_{k+1}}^2 \mathbf{h}} = \sqrt{\mathbf{h}^T \mathbf{B}_d \mathbf{B}_d^T \mathbf{h}} \sigma(\mathbf{z}_k^*, u_k)$. This has the same form as the stability constraint, allowing (20) to be used to transform the state constraint into an SOC constraint with $\bar{\mathbf{A}}_3 = \text{diag}\left\{w_s \sqrt{\gamma_5^*}, 0\right\}$, $\bar{\mathbf{b}}_3 = \left[w_s \frac{\gamma_4^*}{2\sqrt{\gamma_5^*}}, w_s \sqrt{\gamma_3^* - \frac{\gamma_4^{*2}}{4\gamma_5^*}} \right]^T$, $\bar{\mathbf{c}}_3 = [-\mathbf{h}^T \mathbf{B}_d \gamma_2^*, 0]^T$ and $\bar{\mathbf{d}}_3 = -\mathbf{h}^T \mathbf{A}_d \mathbf{z}_k^* - \mathbf{h}^T \mathbf{B}_d \gamma_1^* + b$, with $w_s = \rho(\delta) \sqrt{\mathbf{h}^T \mathbf{B}_d \mathbf{B}_d^T \mathbf{h}}$ ∎
The full controller algorithm is presented in Algorithm 1.

Remark 4.4: The SOC form is maintained when uncertainty in the measured flat state $\hat{\mathbf{z}}_k$ is considered, but is not shown here for brevity.

V. SIMULATION

Our controller was assessed on three tasks using a horizontal 1-D quadrotor. The quadrotor dynamics, as in [9] and [10], are given by $\ddot{x} = \Gamma \sin(\theta) - \gamma \dot{x}$ and $\dot{\theta} = \frac{1}{\tau}(u - \theta)$ where x is the horizontal position, θ is the pitch angle of the quadrotor, and the commanded pitch angle u is the system input. Here, $\Gamma = 10$, $\gamma = 0.3$ and $\tau = 0.2$ are model parameters. This model is flat in the output $y_k = x_k$, with the flat state $\mathbf{z}_k = [x_k, \dot{x}_k, \ddot{x}_k]^T$. All algorithms are run at 50 Hz, use the same gain matrices $\mathbf{Q}$ and r, and horizons.

As shown in Figure 2a, FMPC with the safety filter (FMPC+SOCP) is compared against an MPC using an inaccurate model ($\Gamma = 20$, $\gamma = 0$, and $\tau = 0.05$, chosen to overestimate thrust with no drag), and discrete linear quadratic regulator (DQLR) using the same inaccurate model, and a trained GPMPC where the uncertain dynamics are modelled as a GP inside a robust MPC formulation [16]. Training data was gathered via Latin hypercube sampling of states within the max and min values seen from the reference trajectories. Squared Exponential kernels were used for all GP kernels, trained offline. For this system, $v_k = \dddot{x}_k$, a higher-order derivative of the position, was measured from the simulation. On a real system, this would need to be estimated, which can be hard due to accumulated noise, presenting a limitation of our approach. The root mean squared errors were $0.02\,\text{m}$ and $0.07\,\text{m}$ for FMPC+SOCP and GPMPC, respectively. The average solve time per step was $(0.018 \pm 0.006)\,\text{s}$ for FMPC+SOCP
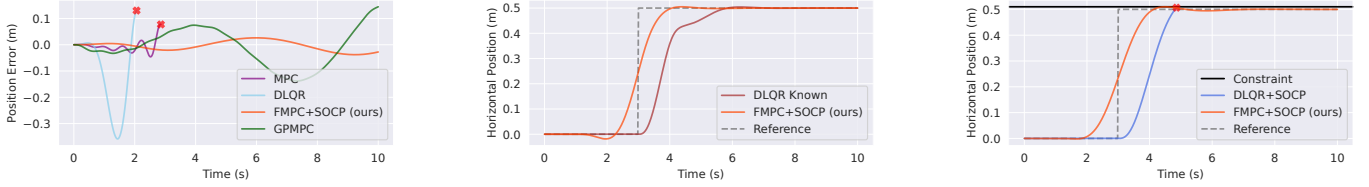
(a) Tracking error for $y_{\text{ref}}(t) = 0.2t\sin(0.9t)$.

(b) Tracking step with $-10° \le u_k \le 10°$.

(c) Step with constraint $x_k \le 0.51$.

Fig. 2. In 2a, the tracking error of our Flat MPC with safety filter (FMPC+SOCP) is compared with an MPC and DLQR with inaccurate prior models, and a trained GPMPC [16]. In 2b the FMPC+SOCP is compared against a DLQR with perfect dynamics knowledge subject to input constraints, and in 2c FMPC+SOCP is compared with DLQR with the safety filter (DLQR+SOCP) subject to a constraint on the position. Red $\times$ indicates a point of controller infeasibility or constraint violation. We see that FPMPC+SOCP performs similarly to GPMPC, while respecting input and state constraints.
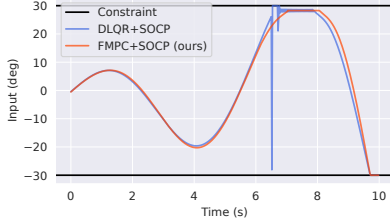

Fig. 3. Input comparison while tracking $y_{\text{ref}}(t) = 0.2t\sin(0.9t)$ subject to velocity constraints $\dot{x}_k \le 1.0$ and input constraints $|u_k| \le 30°$.

and $(0.29 \pm 0.04)\,$s for GPMPC when run on a 16 GB RAM desktop using an AMD 3900xt CPU.

The FMPC+SOCP evidently performs better than GPMPC, while solving the task more than 10 times faster. This occurs in part because having the GP inside the nonlinear optimization can lead to worse local minima—something that nonlinear MPC is already prone to. Additionally, the GP used in GPMPC requires an independent GP for each dimension of the state, meaning roughly five times as many training points were used in GPMPC than in our approach.

In Figure 2b, FMPC+SOCP is compared with DLQR where (1), and thus (4), are known perfectly, but subject to input constraints $|u_k| \le 10°$. The inclusion of DLQR is pertinent as it represents the infinite-horizon optimal solution. In this case, DLQR's inputs are clipped if the desired input exceeds the limit. We see that even in this case, the FMPC+SOCP outperforms the DLQR, as it is designed to account for input constraints.

Finally, FMPC+SOCP is compared against an inaccurate DLQR with the safety filter (DLQR+SOCP) on tracking a step trajectory with a state constraint $x_k \le 0.51$ in Figure 2c and tracking $y_{\text{ref}}(t) = 0.2t\sin(0.9t)$ subject to velocity constraints $\dot{x}_k \le 1.0$ and input constraints $|u| \le 30°$ in Figure 3. In Figure 2c, the predictive nature of the FMPC+SOCP clearly anticipates the step response and settles faster than the DLQR. Additionally, DLQR+SOCP reaches an infeasible state as it overshoots the reference because it could not predict into the future nor account for the state constraint. The FMPC+SOCP avoided violating the constraint boundary due to its predictive nature, remaining feasible. In Figure 3, we see how the FMPC input anticipates the velocity limit, resulting in smooth inputs that never reach their limits. DLQR, however, gives very aggressive, and potentially damaging, changes in input to avoid violations.

VI. Conclusion

Current safe learning-based controllers' real-time applicability has been limited by computational performance. The proposed FMPC+SOCP approach performs similarly to state-of-the-art learning-based controllers, but is 10 times more computationally efficient, while guaranteeing probabilistic asymptotic stability, probabilistic state constraint satisfaction, and input constraint satisfaction.

References

[1] L. Brunke, M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoellig, "Safe learning in robotics: From learning-based control to safe reinforcement learning," *Annu. Rev. Control Rob. Auton. Syst.*, vol. 5, no. 1, pp. 411–444, 2022.

[2] C. J. Ostafew, A. P. Schoellig, and T. D. Barfoot, "Robust constrained learning-based NMPC enabling reliable mobile robot path tracking," *Int. J. Rob. Res.*, vol. 35, no. 13, pp. 1547–1563, 2016.

[3] G. Torrente, E. Kaufmann, P. Fohn, and D. Scaramuzza, "Data-driven MPC for quadrotors," *IEEE Rob. Autom. Lett.*, vol. 6, no. 2, pp. 3769–3776, Apr. 2021.

[4] J. Quinonero-Candela and C. E. Rasmussen, "A unifying view of sparse approximate Gaussian process regression," *J. Machine Learning Research*, vol. 6, pp. 1939–1959, 2005.

[5] M. Fliess, J. Levine, P. Martin, and P. Rouchon, "Flatness and defect of non-linear systems: Introductory theory and examples," *Int. J. Control*, vol. 61, no. 6, pp. 1327–1361, 1995.

[6] M. Greeff and A. P. Schoellig, "Flatness-based model predictive control for quadrotor trajectory tracking," in *2018 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, Madrid, Spaine, 2018.

[7] A. Isidori, *Nonlinear Control Systems*. London: Springer, 1995.

[8] H. Sira-Ramirez and S. K. Agrawal, *Differentially Flat Systems*. Boca Raton: CRC Press, Taylor & Francis Group, 2004.

[9] M. Greeff and A. P. Schoellig, "Exploiting differential flatness for robust learning-based tracking control using gaussian processes," *IEEE Control Syst. Lett.*, vol. 5, no. 4, pp. 1121–1126, 2021.

[10] M. Greeff, A. W. Hall, and A. P. Schoellig, "Learning a stability filter for uncertain differentially flat systems using Gaussian processes," in *2021 60th IEEE Conf. Decision and Control (CDC)*. IEEE, 2021, pp. 789–794.

[11] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA: The MIT Press, 2006.

[12] L. Grune and J. Pannek, *Nonlinear Model Predictive Control: Theory and Algorithms*. Switzerland: Springer, 2017.

[13] F. Borrelli, A. Bemporad, and M. Morari, *Predictive Control for Linear and Hybrid Systems*. Cambridge: Cambridge University Press, 2017.

[14] F. Castaneda, J. J. Choi, B. Zhang, C. J. Tomlin, and K. Sreenath, "Gaussian process-based min-norm stabilizing controller for control-affine systems with uncertain input effects and dynamics," in *2021 American Control Conf. (ACC)*. IEEE, 2021, pp. 3683–3690.

[15] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Information-theoretic regret bounds for Gaussian process optimization in the bandit setting," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3250–3265, 2012.

[16] L. Hewing, J. Kabzan, and M. N. Zeilinger, "Cautious model predictive control using Gaussian process regression," *IEEE Trans. Control Syst. Technol.*, vol. 28, no. 6, pp. 2736–2743, 2020.