# Feedback Linearization using Gaussian Processes

Jonas Umlauft, Thomas Beckers, Melanie Kimmel, Sandra Hirche

*Abstract*— Data-driven approaches from machine learning provide powerful tools to identify dynamical systems with limited prior knowledge of the model structure. This paper utilizes Gaussian processes, a Bayesian nonparametric approach, to learn a model for feedback linearization. By using a proper kernel structure, the training data for identification is collected while an existing controller runs the system. Using the identified dynamics, an improved controller, based on feedback linearization, is proposed. The analysis shows that the resulting system is globally uniformly ultimately bounded. We further derive a relationship between the training data of the system and the size of the ultimate bound to which the system converges with a certain probability. A simulation of a robotic system illustrates the proposed method.

## I. INTRODUCTION

Feedback linearization of nonlinear system allows to control many real-world systems with well-understood techniques from linear control theory [1]. However, feedback linearization requires precise model knowledge to ensure the exact cancellation of all nonlinearities in the system. Therefore, precise identification of the nonlinear dynamical system is a crucial step for applying feedback linearization. However, this becomes more and more challenging, as control is increasingly applied to highly complex systems, especially, when the dynamics cannot be derived from first principles. Since classical system identification uses observations of the system's behavior to select parametric models and tune its parameters [2], it is prone to fail, if no structural knowledge is available. Therefore, data-driven approaches gain popularity because they allow generalization outside of the training data with only a minimum set of prior assumptions [3]. A mathematically very profound method is Bayesian nonparametric modeling which unifies the demanded properties: Incorporate prior knowledge through the Bayesian framework and provide sufficient flexibility through their nonparametric nature.

In this work, we focus on Gaussian processes (GPs) as a Bayesian nonparametric model, which already was successfully applied to the identification of dynamical systems [4], [5]. However, a formal analysis from a control point of view is missing. The necessary convergence analysis of a Gaussian process dynamical model was first conducted in [6], [7] and the work in [8] uses a GP to safely explore the state space of a dynamical system. For the stability analysis, it makes use of a further favorable property of a GP: It provides, along with the estimated function value, a measure for the model fidelity depending on the state space region. Based on the distance to training data, it provides upper bounds for the difference between true and inferred function value as derived in [9]. The knowledge of its own ignorance represents a benefit over many other representation approaches.

To deal with the lack of model knowledge in feedback linearization, the work in [10] focuses on a robust design, under the assumption that the effect of the control input on the state is known. Also [11] considers uncertainties in the given dynamics, but without specifying how these are obtained. The methods in [12], [13] use an adaptive scheme, however, nonparametric identification methods for control affine systems have not been exploited.

As main contribution, we propose the identification with Gaussian processes and show that the system with the proposed feedback linearizing controller is globally uniformly ultimately bounded. For the identification, we transfer the knowledge of the control affine structure of the system into the kernel function. This allows to identify the system in closed-loop, while an arbitrary controller is running the system. Further analysis shows how the ultimate bound is reduced as more knowledge (training data) is available and derives a relationship between training data and size of the set to which the system converges for a given probability.

The remainder of this paper is structured as follows: After defining the considered problem setting in Sec. II, Sec. III introduces the identification method based on Gaussian processes. Section IV focuses on the proposed controller and analyzes its convergence properties. It is followed by a numerical illustration for an example system in Sec. V and a conclusion in Sec. VI.

## II. PROBLEM FORMULATION

Consider a state-feedback linearizable single-input system with the controllable canonical form[1]

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = x_3$$
$$\cdots$$
$$\dot{x}_n = f(\boldsymbol{x}) + g(\boldsymbol{x})u, \quad \boldsymbol{x}_0 = \boldsymbol{x}(0), \qquad (1)$$

All authors are members of the Chair of Information-oriented Control, Department of Electrical and Computer Engineering, Technical University of Munich, D-80333 Munich, Germany [jonas.umlauft, t.beckers, melanie.kimmel, hirche]@tum.de

[1]**Notation:** Lower/upper case bold symbols denote vectors/matrices, respectively. $\boldsymbol{I}_n$ denotes the $n \times n$ identity matrix and $\boldsymbol{A} \succ 0$ the positive definiteness of matrix $\boldsymbol{A}$. $\mathbb{E}[\cdot]$ denotes the expected value, $\mathbb{V}[\cdot]$ the variance of a random variable. $\mathbb{R}_+^0$, $\mathbb{R}_+$ denote all real positive numbers with and without the zero, respectively. The norm $\|\cdot\|$ denotes Euclidean norm. Normal distributed random variables with mean $\mu$ and variance $\sigma$ are denoted by $\mathcal{N}(\mu, \sigma)$. $\boldsymbol{a}_{1:n}$ denotes the first $n$ elements of $\boldsymbol{a}$ and $\sigma_{\max}(\boldsymbol{A}), \sigma_{\min}(\boldsymbol{A})$ the maximal and minimal singular value of a matrix $\boldsymbol{A}$. The $\boldsymbol{x}$ dependencies of $f, \hat{f}, g, \hat{g}$ are partially omitted for notational convenience.

where $\boldsymbol{x} = [x_1 \; x_2 \cdots x_n]^{\mathsf{T}} \in \mathcal{X} \subseteq \mathbb{R}^n$, $u \in \mathcal{U} \subseteq \mathbb{R}$ and the full state vector $\boldsymbol{x}$ is assumed measurable. In addition, the following is assumed:

*Assumption 1:* The unknown functions $f : \mathcal{X} \to \mathbb{R}$, $g : \mathcal{X} \to \mathbb{R}$ are bounded and infinitely differentiable.

*Assumption 2:* The system's relative degree is equal to the system order $n$ for all $\boldsymbol{x} \in \mathcal{X}$.

*Assumption 3:* The sign of $g(\boldsymbol{x})$ is known for at least one $\boldsymbol{x} \in \mathcal{X}$.

Assumptions 1 and 2 are common assumptions for feedback linearization. While the first ensures existence of a unique solution to (1), the second ensures, that there are no internal dynamics. Assumption 3 does not impose any severe restriction, since the sign of $g(\boldsymbol{x})$ can be determined from measurements which are assumed to be available (see below).

From Assumption 2 it holds $g(\boldsymbol{x}) \neq 0$, $\forall \boldsymbol{x} \in \mathcal{X}$ (as $g(\boldsymbol{x}) = 0$ leads to an undefined relative degree). This leads, with continuity from Assumption 1, to a constant sign of $g$. Without loss of generality, we take $g$ to be strictly positive

$$g(\boldsymbol{x}) > 0, \quad \forall \boldsymbol{x} \in \mathcal{X}. \tag{2}$$

The first objective is to approximate the unknown functions $f(\boldsymbol{x})$, $g(\boldsymbol{x})$ through $\hat{f}(\boldsymbol{x})$, $\hat{g}(\boldsymbol{x})$, respectively. Based on these estimates, we aim for feedback linearization using

$$u_f := u = \frac{1}{\hat{g}(\boldsymbol{x})} \left( -\hat{f}(\boldsymbol{x}) + \nu \right), \tag{3}$$

where $\nu \in \mathbb{R}$ is the input to the resulting approximately linearized system.

To determine the estimates $\hat{f}(\boldsymbol{x})$ and $\hat{g}(\boldsymbol{x})$, it is assumed, that a controller $u = u_0(\boldsymbol{x})$ exists, which keeps the trajectories $\boldsymbol{x}(t)$ uniformly bounded for a finite time. Thus, $\exists \; T > 0, \epsilon > 0$ such that $\|\boldsymbol{x}(t)\| < \epsilon, \forall t \in [0, \; T]$. This is necessary to collect $N$ training data pairs of the system (1), without risking damage. Since no assumption on the stability of $f(\boldsymbol{x})$ is made, it is not possible to set $u = 0$ and collect training data of $f(\boldsymbol{x})$ separately from $g(\boldsymbol{x})$. However, since the controller only needs to ensure finite time boundedness, its existence is a very mild assumption. In addition to the state vector $\boldsymbol{x}$, we assume availability of noisy observations $y = \dot{x}_n + \epsilon$, where $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$, resulting in the data set[2]

$$\mathcal{D} = \left\{ \boldsymbol{x}^{(i)}, y^{(i)} \right\}_{i=1}^{N}.$$

Note, that the training points do not need to be samples along a single trajectory but are independent of any time indexing. The general idea is summarized in the following three steps:

- Collect training data $\mathcal{D}$ using the controller $u_0$, e.g., by initializing the system at various $\boldsymbol{x}_0$.
- Perform identification by building estimates $\hat{f}$, $\hat{g}$ using GPs and Assumptions 1-3 as described in Sec. III.

---

[2]For practical applications, time derivatives can be approximated through finite differences. The approximation error is then considered as part of the measurement noise (even though it is not necessarily Gaussian).

- Employ the feedback linearization controller (3) instead of $u_0$ to control the system as described in Sec. IV.

By initially considering the from in (1), we implicitly assume that the states of the canonical form are known.

## III. IDENTIFICATION OF CONTROL AFFINE SYSTEMS USING GAUSSIAN PROCESSES

We aim to identify $f$, $g$ using Gaussian process regression as reviewed in the following.

### A. Gaussian Process Basics

Given noisy measurements of a function $f_{\text{true}} \colon \mathcal{X} \to \mathbb{R}$

$$y_f^{(i)} = f_{\text{true}} \left( \boldsymbol{x}^{(i)} \right) + \epsilon^{(i)}$$

with inputs $\boldsymbol{x}^{(i)} \in \mathcal{X}$, $\epsilon^{(i)} \sim \mathcal{N}(0, \sigma_n^2)$ and $i = 1, \dots, N$. The function is modeled with a stochastic process $f_{\mathcal{GP}}(\boldsymbol{x})$ which assigns a joint Gaussian distribution to any finite subset $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_M\} \subset \mathcal{X}$ in a continuous domain. The Gaussian process, also considered as distribution over functions [14], is denoted by

$$f_{\mathcal{GP}}(\boldsymbol{x}) \sim \mathcal{GP} \left( m(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}') \right).$$

The GP is fully specified by a mean $m(\boldsymbol{x}) : \mathcal{X} \to \mathbb{R}$ and covariance $k(\boldsymbol{x}, \boldsymbol{x}') : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ function. The first is employed to include prior knowledge regarding the unknown function, but is also often set to zero without loss of generality. The latter, also called kernel function, determines among other things the resulting smoothness properties of $f_{\mathcal{GP}}(\boldsymbol{x})$. Both, mean and kernel function, are parameterized by hyperparameters, concatenated in $\boldsymbol{\psi}$. According to Bayesian principles, the optimal hyperparameters to the observed data are obtained by maximizing the likelihood

$$\boldsymbol{\psi}^* = \arg\max_{\boldsymbol{\psi}} \log p(\boldsymbol{y}_f | \boldsymbol{X}, \boldsymbol{\psi}), \tag{4}$$

$$\log p(\boldsymbol{y}_f | \boldsymbol{X}, \boldsymbol{\psi}) = \frac{1}{2} \left( \boldsymbol{y}_f^T \boldsymbol{K} \boldsymbol{y}_f - \log \det \boldsymbol{K} - N \log(2\pi) \right),$$

where

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}^{(1)} & \cdots & \boldsymbol{x}^{(N)} \end{bmatrix} \in \mathbb{R}^{n \times N},$$

$$\boldsymbol{y}_f = \begin{bmatrix} y_f^{(1)} & \cdots & y_f^{(N)} \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^N,$$

concatenate the input/output data, respectively and

$$\boldsymbol{K} = \begin{bmatrix} k\left(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(1)}\right) & \cdots & k\left(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(N)}\right) \\ \vdots & \ddots & \vdots \\ k\left(\boldsymbol{x}^{(N)}, \boldsymbol{x}^{(1)}\right) & \cdots & k\left(\boldsymbol{x}^{(N)}, \boldsymbol{x}^{(N)}\right) \end{bmatrix} \in \mathbb{R}^{N \times N} \tag{5}$$

concatenates pairwise kernel evaluations of the input training data. The optimization is generally non-convex, however it is commonly solved successfully with gradient-based methods [14]. The effect of local minima and the resulting suboptimal identification is discussed in Remark 1.

Gaussian process regression uses the joint Gaussian distribution of training data and a test input $\boldsymbol{x}^*$

$$\begin{bmatrix} f_{\mathcal{GP}}(\boldsymbol{x}^*) \\ \boldsymbol{y}_f \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} m(\boldsymbol{x}^*) \\ \boldsymbol{m_X} \end{bmatrix}, \begin{bmatrix} k^* & \boldsymbol{k}^{\mathsf{T}} \\ \boldsymbol{k} & \boldsymbol{K} + \sigma_n^2 \boldsymbol{I}_N \end{bmatrix} \right),$$

with

$$\boldsymbol{m_X} = \begin{bmatrix} m\left(\boldsymbol{x}^{(1)}\right) & \cdots & m\left(\boldsymbol{x}^{(N)}\right) \end{bmatrix}^{\mathsf{T}},$$

to infer posterior mean and variance through conditioning

$$\mu(\boldsymbol{x}^*) := \mathbb{E}\left[f_{\mathcal{GP}}(\boldsymbol{x}^*)|\mathcal{D}\right] \tag{6}$$
$$= m(\boldsymbol{x}^*) + \boldsymbol{k}^{\mathsf{T}}(\boldsymbol{K} + \sigma_n^2 \boldsymbol{I}_N)^{-1}(\boldsymbol{y}_f - \boldsymbol{m_X}),$$
$$\sigma(\boldsymbol{x}^*) := \mathbb{V}\left[f_{\mathcal{GP}}(\boldsymbol{x}^*)|\mathcal{D}\right] = k^* - \boldsymbol{k}^{\mathsf{T}}(\boldsymbol{K} + \sigma_n^2 \boldsymbol{I}_N)^{-1}\boldsymbol{k}, \tag{7}$$

with

$$k^* = k(\boldsymbol{x}^*, \boldsymbol{x}^*),$$
$$\boldsymbol{k} = \begin{bmatrix} k\left(\boldsymbol{x}^{(1)}, \boldsymbol{x}^*\right) & \cdots & k\left(\boldsymbol{x}^{(N)}, \boldsymbol{x}^*\right) \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^N. \tag{8}$$

Going back to the original problem, the difficulty here is that individual observations of the function $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ from (1) are not available in $\mathcal{D}$. Since $\dot{x}_n = f(\boldsymbol{x})$ is possibly unstable, we must run the system in closed-loop to take measurements. Thus, we aim to identify the functions $f, g$ from only observing their sum, as given in the data set $\mathcal{D}$. Therefore, the knowledge regarding the control affine structure of (1) will be exploited as follows.

### B. Expressing Structure in Kernels

The kernel function of a GP does not only express its smoothness properties but is also used to express prior knowledge on the structure of the function [15]. Note that this it not equivalent to assuming a parametric structure of the function as in classical system identification. The kernel encodes that the function is represented as a sum or a product but leaves each function unlimited flexibility.

*1) Sum of functions:* Suppose $f_a, f_b \colon \mathcal{X} \to \mathbb{R}$ are drawn independently from two GP priors

$$f_a(\boldsymbol{x}) \sim \mathcal{GP}\left(m_a(\boldsymbol{x}), k_a(\boldsymbol{x}, \boldsymbol{x}')\right),$$
$$f_b(\boldsymbol{x}) \sim \mathcal{GP}\left(m_b(\boldsymbol{x}), k_b(\boldsymbol{x}, \boldsymbol{x}')\right),$$

and are composed to $f_{\text{sum}}(\boldsymbol{x}) = f_a(\boldsymbol{x}) + f_b(\boldsymbol{x})$ then

$$f_{\text{sum}}(\boldsymbol{x}) \sim \mathcal{GP}\left(m_a(\boldsymbol{x}) + m_b(\boldsymbol{x}), k_a(\boldsymbol{x}, \boldsymbol{x}') + k_b(\boldsymbol{x}, \boldsymbol{x}')\right)$$

is another GP. Assuming that there are measurements

$$y_{\text{sum}}^{(i)} = f_{\text{sum}}\left(\boldsymbol{x}^{(i)}\right) + \epsilon^{(i)} = f_a\left(\boldsymbol{x}^{(i)}\right) + f_b\left(\boldsymbol{x}^{(i)}\right) + \epsilon^{(i)},$$

of the sum of the functions available with $\epsilon^{(i)} \sim \mathcal{N}(0, \sigma_n^2)$ and $i = 1, \ldots, N$. Then, the joint distribution ($m_a(\boldsymbol{x}) = m_b(\boldsymbol{x}) = 0$ for notational simplicity)

$$\begin{bmatrix} f_a(\boldsymbol{x}^*) \\ f_b(\boldsymbol{x}^*) \\ \boldsymbol{y}_{\text{sum}} \end{bmatrix} \sim \mathcal{N}\left(\boldsymbol{0}, \begin{bmatrix} k_a^* & \boldsymbol{0} & \boldsymbol{k}_a^{\mathsf{T}} \\ \boldsymbol{0} & k_b^* & \boldsymbol{k}_b^{\mathsf{T}} \\ \boldsymbol{k}_a & \boldsymbol{k}_b & \boldsymbol{K}_a + \boldsymbol{K}_b + \sigma_n^2 \boldsymbol{I}_N \end{bmatrix}\right)$$

with $\boldsymbol{k}_a, \boldsymbol{k}_b, k_a^*, k_b^*$ defined according to (8), is used for conditioning on the output of the individual functions

$$f_a(\boldsymbol{x}^*)|\boldsymbol{y}_{\text{sum}} \sim \mathcal{N}\left(\boldsymbol{k}_a^{\mathsf{T}} \boldsymbol{K}_{\text{sum}}^{-1} \boldsymbol{y}_{\text{sum}}, k_a^* - \boldsymbol{k}_a^{\mathsf{T}} \boldsymbol{K}_{\text{sum}}^{-1} \boldsymbol{k}_a\right),$$
$$f_b(\boldsymbol{x}^*)|\boldsymbol{y}_{\text{sum}} \sim \mathcal{N}\left(\boldsymbol{k}_b^{\mathsf{T}} \boldsymbol{K}_{\text{sum}}^{-1} \boldsymbol{y}_{\text{sum}}, k_b^* - \boldsymbol{k}_b^{\mathsf{T}} \boldsymbol{K}_{\text{sum}}^{-1} \boldsymbol{k}_b\right),$$

where $\boldsymbol{K}_{\text{sum}} = \boldsymbol{K}_a + \boldsymbol{K}_b + \sigma_n^2 \boldsymbol{I}_N$ with $\boldsymbol{K}_a, \boldsymbol{K}_b$ according to (5). The hyperparameter vector is extended accordingly, $\boldsymbol{\psi}_{\text{sum}} = [\boldsymbol{\psi}_a^{\mathsf{T}} \; \boldsymbol{\psi}_b^{\mathsf{T}}]^{\mathsf{T}}$ and also determined by optimization (4),

where $\boldsymbol{K} = \boldsymbol{K}_{\text{sum}}$ and $\boldsymbol{y}_f = \boldsymbol{y}_{\text{sum}}$. This enables us to infer the posterior mean and variance functions of $f_a, f_b$ separately, using only observations of their sum.

*2) Product with known function:* The knowledge that an unknown function $f_h(\boldsymbol{x}) \colon \mathcal{X} \to \mathbb{R}$ is multiplied with a known function $h(\boldsymbol{x}) \colon \mathcal{X} \to \mathbb{R}$, can also be represented in the kernel of a GP: Suppose measurements of the product

$$y_{\text{prod}}^{(i)} = f_{\text{prod}}\left(\boldsymbol{x}^{(i)}\right) + \epsilon^{(i)} = f_h\left(\boldsymbol{x}^{(i)}\right) h\left(\boldsymbol{x}^{(i)}\right) + \epsilon^{(i)},$$

with $\epsilon^{(i)} \sim \mathcal{N}(0, \sigma_n^2)$, $i = 1, \ldots, N$ are taken and $f_h \sim \mathcal{GP}(0, k_h(\boldsymbol{x}, \boldsymbol{x}'))$, then $f_{\text{prod}}(\boldsymbol{x})$ is also a GP with scaled kernel ($m_h(\boldsymbol{x}) = 0$ for notational simplicity)

$$f_{\text{prod}}(\boldsymbol{x}) \sim \mathcal{GP}(0, h(\boldsymbol{x})k_h(\boldsymbol{x}, \boldsymbol{x}')h(\boldsymbol{x}')).$$

The resulting joint distribution is

$$\begin{bmatrix} f_h(\boldsymbol{x}^*) \\ \boldsymbol{y}_{\text{prod}} \end{bmatrix} \sim \mathcal{N}\left(\boldsymbol{0}, \begin{bmatrix} h^* k_h^* h^* & h^* \boldsymbol{k}_h^{\mathsf{T}} \boldsymbol{H}^{\mathsf{T}} \\ \boldsymbol{H} \boldsymbol{k}_h h^* & \boldsymbol{H}^{\mathsf{T}} \boldsymbol{K}_h \boldsymbol{H} + \sigma_n^2 \boldsymbol{I}_N \end{bmatrix}\right),$$

where $\boldsymbol{H} = \text{diag}\left(h\left(\boldsymbol{x}^{(1)}\right), \ldots, h\left(\boldsymbol{x}^{(N)}\right)\right) \in \mathbb{R}^{N \times N}$, $h^* = h(\boldsymbol{x}^*)$ and $\boldsymbol{k}_h$, $k_h^*$, $\boldsymbol{K}_h$ as in (8) and (5), respectively. Defining $\boldsymbol{K}_{\text{prod}} = \boldsymbol{H}^{\mathsf{T}} \boldsymbol{K}_h \boldsymbol{H} + \sigma_n^2 \boldsymbol{I}_N$ and conditioning leads to inference for function $f_h$

$$f_h(\boldsymbol{x}^*)|\boldsymbol{y}_{\text{prod}} \sim \mathcal{N}\big(h^* \boldsymbol{k}_h^{\mathsf{T}} \boldsymbol{H}^{\mathsf{T}} \boldsymbol{K}_{\text{prod}}^{-1} \boldsymbol{y}_{\text{prod}},$$
$$h^* k_h^* h^* - h^* \boldsymbol{k}_h^{\mathsf{T}} \boldsymbol{H}^{\mathsf{T}} \boldsymbol{K}_{\text{prod}}^{-1} \boldsymbol{H} \boldsymbol{k}_h h^*\big).$$

*3) Restriction to positive domain:* The true function $g(\boldsymbol{x})$ is strictly positive $\forall \boldsymbol{x} \in \mathcal{X}$ and therefore this must also hold for its estimate $\hat{g}(\boldsymbol{x})$ to obtain bounded control signals in (3). This is ensured as follows: According to [6], the posterior mean function $\mu(\boldsymbol{x})$ for a squared exponential kernel is bounded. Thus, a constant $b_\mu > 0$ exists such that

$$\mu_{m=0}(\boldsymbol{x}) + b_\mu > 0, \quad \forall \boldsymbol{x} \in \mathcal{X},$$

where $\mu_{m=0}$ denotes the posterior mean function with zero prior mean. There exist methods to compute arbitrarily tight bounds for $\mu_{m=0}(\boldsymbol{x})$ on a compact set [16]. Thus, without elaborating further details, we assume a $b_\mu$ is known, which is set as constant prior mean, for which $\mu(\boldsymbol{x}) > 0$, $\forall \boldsymbol{x} \in \mathcal{X}$.

### C. GP Identification for Control Affine Systems

In the following, we bring all mentioned techniques together to identify control affine system $f(\boldsymbol{x}) + g(\boldsymbol{x})u(\boldsymbol{x})$ using GPs. We set the prior mean functions for $f$ and $g$ to

$$m_f(\boldsymbol{x}) = 0, \quad m_g(\boldsymbol{x}) = b_\mu, \tag{9}$$

and use a composite of squared exponential (SE) kernels with automatic relevance determination

$$k(\boldsymbol{x}, \boldsymbol{x}') = k_f(\boldsymbol{x}, \boldsymbol{x}') + u_0(\boldsymbol{x})k_g(\boldsymbol{x}, \boldsymbol{x}')u_0(\boldsymbol{x}'). \tag{10}$$

with

$$k_f(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2 \exp\left(\sum_{j=1}^n \frac{(x_j - x_j')^2}{-2l_{j,f}^2}\right),$$

$$k_g(\boldsymbol{x}, \boldsymbol{x}') = \sigma_g^2 \exp\left(\sum_{j=1}^n \frac{(x_j - x_j')^2}{-2l_{j,g}^2}\right),$$

where the hyperparameters are lengthscales $l_{j,f}, l_{j,g} \in \mathbb{R}_+$, $j = 1, \ldots, n$, signal variance $\sigma_f, \sigma_g \in \mathbb{R}_+^0$ and observation noise $\sigma_n \in \mathbb{R}_+^0$. They are concatenated in

$$\boldsymbol{\psi}_{gf} = [l_{1,f} \; l_{1,g} \; \cdots \; l_{n,f} \; l_{n,g} \; \sigma_f^2 \; \sigma_g^2 \; \sigma_n^2]^\mathsf{T},$$

and optimized using (4). The joint distribution with test input $\boldsymbol{x}^*$ is

$$\begin{bmatrix} f(\boldsymbol{x}^*) \\ g(\boldsymbol{x}^*) \\ \boldsymbol{y} \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} 0 \\ b_\mu \\ \boldsymbol{U}_0 \boldsymbol{b_\mu} \end{bmatrix}, \begin{bmatrix} k_f^* & 0 & \boldsymbol{k}_f^\mathsf{T} \\ 0 & k_g^* & \boldsymbol{k}_g^\mathsf{T} \boldsymbol{U}_0^\mathsf{T} u_0^* \\ \boldsymbol{k}_f & \boldsymbol{U}_0 \boldsymbol{k}_g u_0^* & \boldsymbol{K}_{fg} \end{bmatrix} \right),$$

where $\boldsymbol{U}_0 = \text{diag} \left( u_0 \left( \boldsymbol{x}^{(1)} \right), \ldots, u_0 \left( \boldsymbol{x}^{(N)} \right) \right) \in \mathbb{R}^{N \times N}$, $u_0^* = u_0(\boldsymbol{x}^*)$, $\boldsymbol{b_\mu} = [b_\mu \; \cdots \; b_\mu]^\mathsf{T} \in \mathbb{R}^N$,

$$\boldsymbol{K}_{fg} = \boldsymbol{K}_f + \boldsymbol{U}_0^\mathsf{T} \boldsymbol{K}_g \boldsymbol{U}_0 + \sigma_n^2 \boldsymbol{I}_n$$

and $\boldsymbol{k}_f, \boldsymbol{k}_g, \boldsymbol{K}_f, \boldsymbol{K}_g$ are build using $k_f(\boldsymbol{x}, \boldsymbol{x}'), k_g(\boldsymbol{x}, \boldsymbol{x}')$ analog to (8) and (5). The training output

$$\boldsymbol{y} = \begin{bmatrix} y^{(1)} & \cdots & y^{(N)} \end{bmatrix}^\mathsf{T} \in \mathbb{R}^N$$

and input values $\boldsymbol{x}^{(i)}$, $i = 1, \ldots, N$ are given in the data set $\mathcal{D}$. The resulting posterior mean function of the two functions are set to the estimates $\hat{f}(\boldsymbol{x}), \hat{g}(\boldsymbol{x})$ given by

$$\hat{f}(\boldsymbol{x}^*) := \mu_f(\boldsymbol{x}^*) = \boldsymbol{k}_f^\mathsf{T} \boldsymbol{K}_{fg}^{-1} (\boldsymbol{y} - \boldsymbol{U}_0 \boldsymbol{b_\mu}) \qquad (11)$$

$$\hat{g}(\boldsymbol{x}^*) := \mu_g(\boldsymbol{x}^*) = b_\mu + u_0^* \boldsymbol{k}_g^\mathsf{T} \boldsymbol{U}_0 \boldsymbol{K}_{fg}^{-1} (\boldsymbol{y} - \boldsymbol{U}_0 \boldsymbol{b_\mu}).$$

*Lemma 1:* Consider a control affine system (1) with Assumptions 1-3 and corresponding training data in $\mathcal{D}$. Then the estimates $\hat{f}(\boldsymbol{x})$ and $\hat{g}(\boldsymbol{x})$ represented by a Gaussian process with SE kernels (10) and prior mean (9) are bounded, infinitely differentiable and $\hat{g}(\boldsymbol{x}) > 0$ holds $\forall \boldsymbol{x} \in \mathcal{X}$.

*Proof:* The differentiability and boundedness follows from the SE kernel according to [14] and [6], respectively. The strict positivity of $\hat{g}(\boldsymbol{x})$ follows from the boundedness and the choice of the constant prior mean in (9). ∎

*Remark 1:* Obviously, there exist infinitely many possibilities to find two functions which add up to the same function. Therefore, by only using observations of the sum, learning the unique correct individual functions cannot be expected. The estimates in (11) are only one of many possibilities (selected by the hyperparameters) which are in line with the training data. However, additional knowledge regarding the structure of $f(\boldsymbol{x}), g(\boldsymbol{x})$, e.g., dependence only a subset of the state variables, can be transfered into the kernel: By setting the lengthscales $l_{j,f} = l_{j',g} = \infty$ for the states $j, j'$ on which they do not depend, it is ensured, that also the estimates show no dependency on these input directions. This simplifies the optimization of hyperparameters since it takes place in a lower dimensional space. In the general case, the identification possibly fails in separating the contribution of $f$ and $g$. The limits have not been studied well. Nevertheless, this does not affect Lemma 1 and only compromises performance but not the following results on convergence.

## IV. CONVERGENCE ANALYSIS FOR FEEDBACK LINEARIZING CONTROL

This section proposes the feedback linearizing controller based on estimates $\hat{f}, \hat{g}$ from (11) and analyzes the convergence of the closed loop system.
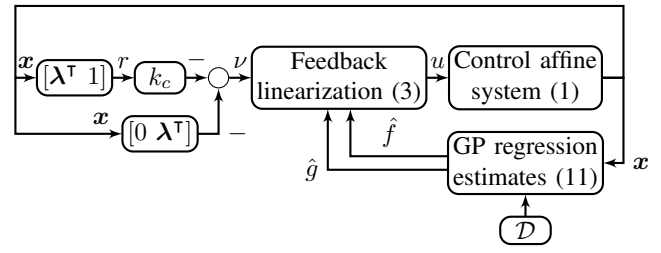


Fig. 1. The overall design of the framework

### A. Control Law

Consider the filtered scalar state $r \in \mathbb{R}$, defined as

$$r = \begin{bmatrix} \boldsymbol{\lambda}^\mathsf{T} & 1 \end{bmatrix} \boldsymbol{x}, \qquad (12)$$

where $\boldsymbol{\lambda} = [\lambda_1 \; \lambda_2 \cdots \lambda_{n-1}]^\mathsf{T} \in \mathbb{R}^{n-1}$ is a coefficient vector such that for $s \in \mathbb{C}$ the polynomial $s^{n-1} + \lambda_{n-1} s^{n-2} + \cdots + \lambda_1$ is Hurwitz. Under this condition, the state converges exponentially $\boldsymbol{x} \to \boldsymbol{0}$ as $r \to 0$ [12]. The dynamic of the filtered state is

$$\dot{r} = f(\boldsymbol{x}) + g(\boldsymbol{x}) u_f(\boldsymbol{x}) + \boldsymbol{\lambda}^\mathsf{T} \boldsymbol{x}_{2:n},$$

where $\boldsymbol{x}_{2:n} = [x_2 \cdots x_n]^\mathsf{T} \in \mathbb{R}^{n-1}$ and

$$u_f(\boldsymbol{x}) = \frac{1}{\hat{g}(\boldsymbol{x})} \left( -\hat{f}(\boldsymbol{x}) - k_c r - \boldsymbol{\lambda}^\mathsf{T} \boldsymbol{x}_{2:n} \right), \qquad (13)$$

according to (3) where

$$\nu = -k_c r - \boldsymbol{\lambda}^\mathsf{T} \boldsymbol{x}_{2:n},$$

with $k_c \in \mathbb{R}_+$ is used. The control scheme is shown in Fig. 1.

### B. Convergence Analysis

The closed-loop behavior is analyzed in the following.

*Proposition 1:* There exists a $k_c^* > 0$ such that for every $k_c \geq k_c^*$ the system (1) under Assumptions 1-3 with estimates (11) under control law (13) is globally uniformly ultimately bounded.

*Proof:* Consider the Lyapunov candidate

$$V(\boldsymbol{x}) = r^2/2 \qquad (14)$$

with its time derivative

$$\begin{aligned} \dot{V}(\boldsymbol{x}) &= r\dot{r} = r \left( f + gu + \boldsymbol{\lambda}^\mathsf{T} \boldsymbol{x}_{2:n} \right) \\ &= r \left( f - \bar{g}\hat{f} \right) - k_c \bar{g} r^2 + (1 - \bar{g}) r \boldsymbol{\lambda}^\mathsf{T} \boldsymbol{x}_{2:n}, \quad (15) \end{aligned}$$

where $\bar{g} := g/\hat{g}$ is positive and bounded $\forall \boldsymbol{x} \in \mathcal{X}$ from Lemma 1 and Assumptions 1-3. In addition, define constants $\boldsymbol{b} \in \mathbb{R}^n$, $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n \times n}$, with $\boldsymbol{A}, \boldsymbol{B} \succ 0$ for which

$$\boldsymbol{b}^\mathsf{T} \boldsymbol{x} \geq r \left( f - \bar{g}\hat{f} \right), \qquad \boldsymbol{x}^\mathsf{T} \boldsymbol{A} \boldsymbol{x} \leq \bar{g} r^2,$$

$$\boldsymbol{x}^\mathsf{T} \boldsymbol{B} \boldsymbol{x} \geq (1 - \bar{g}) r \boldsymbol{\lambda}^\mathsf{T} \boldsymbol{x}_{2:n}, \qquad \forall \boldsymbol{x} \in \mathcal{X},$$

holds which exist since $f, \hat{f}, \bar{g}$ are bounded. Therefore,

$$\dot{V}(\boldsymbol{x}) \leq \|\boldsymbol{b}\| \|\boldsymbol{x}\| - k_c \sigma_{\min}(\boldsymbol{A}) \|\boldsymbol{x}\|^2 + \sigma_{\max}(\boldsymbol{B}) \|\boldsymbol{x}\|^2,$$

and there exists a $k_c^* > 0$ such that

$$\sigma_{\max}(\boldsymbol{B}) - k_c^* \sigma_{\min}(\boldsymbol{A}) < 0. \tag{16}$$

For every $k_c > k_c^*$, the Lyapunov function decreases

$$\dot{V}(\boldsymbol{x}) < 0, \quad \forall \boldsymbol{x} \in \mathcal{X} \setminus \mathcal{B}_1$$

outside of the hyperball

$$\mathcal{B}_1 = \left\{ \boldsymbol{x} \in \mathcal{X} \,\middle|\, \|\boldsymbol{x}\| \leq \frac{\|\boldsymbol{b}\|}{k_c \sigma_{\min}(\boldsymbol{A}) - \sigma_{\max}(\boldsymbol{B})} \right\}.$$

The attributes uniform and global hold, since the closed-loop system is time independent and the Lyapunov function is radially unbounded [1]. Therefore, trajectories of the closed-loop system converge for any $\boldsymbol{x}_0 \in \mathbb{R}^n$ to $\mathcal{B}_1$. ∎

*Remark 2:* Since no assumption regarding the available training data are made for Proposition 1, the result holds also when only applying the prior estimates, i.e. the prior mean functions $\hat{f}(\boldsymbol{x}) = m_f(\boldsymbol{x}) = 0$ and $\hat{g}(\boldsymbol{x}) = m_g(\boldsymbol{x}) = b_\mu$ without any training. We also have not assumed anything regarding the precision of $\hat{f}, \hat{g}$, thus the statements hold regardless of a successful hyperparameter optimization as mentioned in Remark 1.

*C. Tighter Bound with Training Data*

It is expected, that the set $\mathcal{B}_1$, to which the trajectories converge according to Proposition 1, is getting smaller as a higher quantity (amount of training data) and quality (distribution of input locations in the training data over $\mathcal{X}$) of training data is available. This relationship is investigated in the following. To keep the analysis tractable, the following assumption is made:

*Assumption 4:* The function $g(\boldsymbol{x})$ is known, thus $\hat{g}(\boldsymbol{x}) = g(\boldsymbol{x})$ and noisy measurements for $f(\boldsymbol{x})$,

$$y_f^{(i)} = f\left(\boldsymbol{x}^{(i)}\right) + \epsilon^{(i)} = \dot{x}_n^{(i)} - g\left(\boldsymbol{x}^{(i)}\right) u_0\left(\boldsymbol{x}^{(i)}\right) + \epsilon^{(i)}$$

with $\epsilon^{(i)} \sim \mathcal{N}(0, \sigma_n^2)$, $i = 1, \ldots, N$ are available.
This assumption is commonly applied for robust feedback linearization, e.g., [10]. Analog to (11), the unknown function is now estimated with

$$\hat{f}(\boldsymbol{x}) := \mu(\boldsymbol{x}) = \boldsymbol{k}_f^\mathsf{T}(\boldsymbol{K}_f + \sigma_n^2 \boldsymbol{I}_N)^{-1} \boldsymbol{y}_f, \tag{17}$$

where $\boldsymbol{y}_f = \left[ y_f^{(1)} \cdots y_f^{(N)} \right]^\mathsf{T}$ and $\boldsymbol{k}_f$, $\boldsymbol{K}_f$ are computed according to (8) and (5) for the SE kernel. Under this assumption, Proposition 1 can be relaxed with respect to the choice of the gain $k_c$.

*Corollary 1:* The system (1) under Assumptions 1-4 with estimate (17) for $f(\boldsymbol{x})$ under control law (13) is globally uniformly ultimately bounded for any $k_c > 0$.

*Proof:* As a special case of Proposition 1, the condition (16) is dropped since $1 - \bar{g} = 0$. This allows to extend the statement for all gains $k_c$. ∎

In the following, we aim to derive a tight bound regarding the precision of the estimate $\hat{f}(\boldsymbol{x})$ using the variance function $\sigma(\boldsymbol{x}): \mathcal{X} \to \mathbb{R}_+^0$ of the GP as defined in (7). It is an indicator for the model fidelity and (for a SE kernel) it is also a measure for the distance to training data.

*Lemma 2:* The variance function of a GP, as defined in (7), is bounded for a SE kernel by

$$0 \leq \sigma(\boldsymbol{x}) \leq \sigma_f^2, \quad \forall \boldsymbol{x} \in \mathcal{X}.$$

*Proof:* Since the SE kernel is bounded by $0 \leq k_f(\boldsymbol{x}, \boldsymbol{x}') \leq \sigma_f^2$ [7], it follows from $\boldsymbol{K}_f \succ 0$. ∎

Since the GP is a probabilistic model in nature, we cannot expect any deterministic statements regarding the error of the estimate $|f - \hat{f}|$, however, according to [9], it is possible to make high probability statements regarding the distance from the true function $f(\boldsymbol{x})$ to the mean function $\mu(\boldsymbol{x})$ using the variance function $\sigma(\boldsymbol{x})$. This requires additional assumptions regarding the complexity of $f(\boldsymbol{x})$.

*Assumption 5:* $f(\boldsymbol{x})$ has a bounded reproducing kernel Hilbert space (RKHS) norm with respect to a continuously differentiable bounded kernel $k$, denoted by $\|f(\boldsymbol{x})\|_k^2 \leq B_f$.

*Lemma 3:* [9] Suppose Assumption 5 holds, then

$$\Pr\left\{ \forall \boldsymbol{x} \in \tilde{\mathcal{X}}, |\mu_N(\boldsymbol{x}) - f(\boldsymbol{x})| \leq \beta_{N+1} \sigma_N(\boldsymbol{x}) \right\} \geq 1 - \delta,$$

holds on a compact set $\tilde{\mathcal{X}} \subset \mathbb{R}^n$, where $\delta \in (0, 1)$, $\beta_{N+1} = \sqrt{2B_f + 300\gamma_{N+1} \log^3((N+1)/\delta)}$ and $\gamma_{N+1}$ is the maximum mutual information that can be obtained about $f(\boldsymbol{x})$ from $N+1$ noisy samples $\boldsymbol{x}_{(1)}, \ldots, \boldsymbol{x}_{(N+1)}$ and $\mu_N(\boldsymbol{x})$ and $\sigma_N(\boldsymbol{x})$ are posterior mean and variance function of a GP for $N$ data points as defined in (6), (7), respectively.

*Proposition 2:* For system (1) under Assumptions 1-5 with estimate (17) for $f(\boldsymbol{x})$ based on $N$ training points and under control law (13), with probability $1 - \delta$, $\delta \in (0, 1)$, the closed-loop system is globally uniformly ultimately bounded for any $k_c > 0$ to the set

$$\mathcal{B}_2 = \left\{ \boldsymbol{x} \in \mathcal{X} \,\middle|\, \|\boldsymbol{x}\| \leq \frac{\beta_{N+1} \bar{\sigma}}{k_c \left\| \left[ \boldsymbol{\lambda}^\mathsf{T} \ 1 \right] \right\|} \right\},$$

where $\bar{\sigma}$ is an upper bound such that $\bar{\sigma} \geq \sigma(\boldsymbol{x})$ holds $\forall \boldsymbol{x} \in \mathcal{B}_1$ and $\beta_{N+1}$ is defined in Lemma 3.

*Proof:* Proposition 1 guarantees the convergence to a compact set $\mathcal{B}_1 = \tilde{\mathcal{X}}$. Thus, the results from Lemma 3 are applicable and the further analysis can be limited to $\mathcal{B}_1$. Using Lyapunov candidate (14), its time derivative (15) for the case $g - \hat{g} = 0$ (Assumption 4) is

$$\dot{V}(\boldsymbol{x}) = r(f - \hat{f}) - k_c r^2. \tag{18}$$

According to Lemma 3,

$$|f - \hat{f}| \leq \beta_{N+1} \bar{\sigma} \quad \Rightarrow \quad \dot{V}(\boldsymbol{x}) < 0 \quad \forall \boldsymbol{x} \in \mathcal{X} \setminus \mathcal{B}_2,$$

holds with a probability $1 - \delta$, which ensures convergence to a set of radius $\frac{\beta_{N+1} \bar{\sigma}}{k_c}$. Because $r$ converges to $\mathcal{B}_2$, $\boldsymbol{x}$ is ultimately bounded. The attributes uniform and global hold, since the closed-loop system is time independent and the Lyapunov function is radially unbounded [1]. ∎

*Remark 3:* Even though $\mathcal{B}_1$, $\mathcal{B}_2$ from Propositions 1 and 2 can be made arbitrarily small through larger choices of $k_c$, this is undesired due to the resulting high gain controller. Given a fixed set of training data, the convergence set $\mathcal{B}_2$ can be computed from Proposition 2 for any probability $1 - \delta$

(and also vice versa). If the pair of achieved probability $1-\delta$ and corresponding set size $\mathcal{B}_2$ does not meet the required performance, the system must be trained with additional data to reduce $\bar{\sigma}$. The following lemma exploits the behavior of $\sigma(\boldsymbol{x})$ as more training data is added.

*Lemma 4:* For the posterior variance function with SE kernel defined in (7), the following two statements hold:

i) Let $\sigma_N(\boldsymbol{x}), \sigma_{N+1}(\boldsymbol{x})$ denote the variance functions for $N$, $N+1$ training data points, respectively, then

$$\sigma_N(\boldsymbol{x}) > \sigma_{N+1}(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \mathcal{X},$$

ii) If to the existing training points $\boldsymbol{x}^{(i)}$, $i = 1, \ldots, N_0$, training points at the test point $\boldsymbol{x}^* = \boldsymbol{x}^{(i)}$, $i = N_0 + 1, \ldots, N$ are added, then

$$\lim_{N \to \infty} \sigma_N(\boldsymbol{x}^*) = 0 \quad \text{with} \quad \sigma_N = \mathcal{O}(1/N).$$

*Proof:* Statement i) follows directly from constant $k_f^* = \sigma_f^2$ in (7) and $\boldsymbol{K}_f \succ 0$. For statement ii), consider

$$
\begin{aligned}
\sigma_N(\boldsymbol{x}^*) &= \sigma_f^2 - \boldsymbol{k}_f^\mathsf{T}(\boldsymbol{K}_f + \sigma_n^2 \boldsymbol{I})^{-1}\boldsymbol{k}_f \\
&\leq \sigma_f^2 - \|\boldsymbol{k}_f\|^2 \sigma_{\min}\left((\boldsymbol{K}_f + \sigma_n^2 \boldsymbol{I})^{-1}\right) \\
&\leq \sigma_f^2 - \frac{\sum_{i=1}^N k_f(\boldsymbol{x}^*, \boldsymbol{x}^{(i)})^2}{\sigma_{\max}(\boldsymbol{K}_f) + \sigma_n^2} \\
&\leq \sigma_f^2 - \frac{\sum_{i=1}^{N_0} k_f(\boldsymbol{x}^*, \boldsymbol{x}^{(i)})^2 + (N - N_0)\sigma_f^2}{N\sigma_f^2 + \sigma_n^2},
\end{aligned}
$$

where the last inequality is obtained from the Gershgorin circle theorem and the fact $K_{ij} \leq \sigma_f^2$, $\forall i, j = 1, \ldots, N$. Since $\sum_{i=1}^{N_0} k_f(\boldsymbol{x}^*, \boldsymbol{x}^{(i)})^2/\sigma_f^2 - N_0 \sigma_f^2 := k_0$ is constant

$$\lim_{N \to \infty} \sigma_N(\boldsymbol{x}^*) \leq \lim_{N \to \infty} \sigma_f^2\left(1 - \frac{k_0 + N}{N + \sigma_n^2/\sigma_f^2}\right) = 0,$$

which concludes the proof. ∎

Assuming now, that the location of the training points can be chosen arbitrarily in $\mathcal{B}_1$, it is shown in [9], that the maximum mutual information $\gamma_{N+1}$ has a sub-linear dependency on the number of training points $N$ for the SE kernel $\gamma_{N+1} = \mathcal{O}(\log(N)^{n+1})$. Therefore, even though $\beta_{N+1}$ increases with $N$, the product $\beta_N \bar{\sigma}$ decreases as more appropriately chosen points are in the training set.

## V. NUMERICAL ILLUSTRATION

To validate the proposed approach, we conduct a simulation for the following system with $n = 2$ and

$$
\begin{aligned}
\dot{x}_1 &= x_2, \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (19) \\
\dot{x}_2 &= \underbrace{1 - \sin(x_1) + s(x_2)}_{=f(\boldsymbol{x})} + \underbrace{\left(1 + \frac{1}{2}\sin(x_2/2)\right)}_{=g(\boldsymbol{x})} u,
\end{aligned}
$$

where $s(x) = \frac{0.5}{1+\exp(-x/10)}$ is the sigmoidal function. The system is an adapted pendulum system and fulfills Assumptions 1-3. Note, that for the simulation, we are not making use of Assumptions 4 and 5, since we aim to validate the simultaneous identification of $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$.

| $k_c$ | $\lambda$ | $\sigma_n^2$ | $k_p$ | $k_d$ | $\boldsymbol{x}_0$ | $b_\mu$ | $N_{\text{full}}$ | $N_{\text{frac}}$ |
|-------|-----------|--------------|-------|-------|---------------------|---------|-------------------|-------------------|
| 0.3 | 1 | 0.1 | 5 | 2 | $[-3\ 0]^\mathsf{T}$ | 2 | 150 | 6 |

TABLE I

PARAMETERS EMPLOYED IN THE SIMULATION
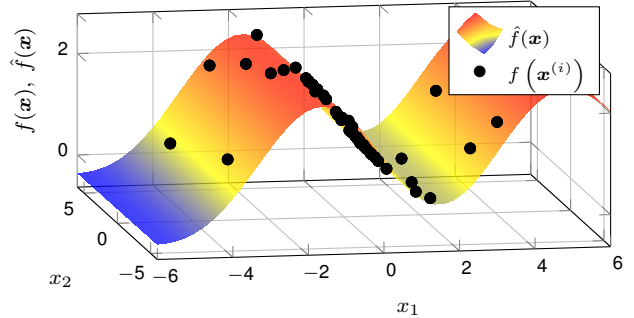


Fig. 2. The estimate $\hat{f}(\boldsymbol{x})$ (surface) along with the test points $f\left(\boldsymbol{x}^{(i)}\right)$, for $\mathcal{D}_{\text{full}}$ (black marks). Consider that $f\left(\boldsymbol{x}^{(i)}\right)$ are not available as measurements for identification but are only added here as reference.

### A. Setup

As described in Sec. II, the simulation is partitioned in the following three steps:

**Collecting training data:** A PD controller

$$u_0(\boldsymbol{x}) = -k_p x_1 - k_d x_2,$$

with parameters from Table I is employed. The system is initialized 3 times at 10 different starting points which are randomly selected from the interval $[-5, 5] \times [-5, 5]$ according to a uniform distribution. Each of the 30 trajectories runs for $T_{\text{train}} = 2s$ and is sampled at time steps of $0.5s$, resulting in $N_{\text{full}} = 150$ training points. This data set is denoted as $\mathcal{D}_{\text{full}}$. For comparison, we also run the identification on a small fraction of the data set $\mathcal{D}_{\text{frac}}$. The observations of $\dot{x}_2$ are perturbed with $\mathcal{N}(0, \sigma_n^2)$ and recorded along with the system state $\boldsymbol{x} = [x_1\ x_2]^\mathsf{T}$.

**Identification:** The likelihood optimization (4) is performed using the optimizer from [14] and initial values $\psi_{gf} = [3\ 3\ 5\ 3\ 3\ 5\ 0.1]^\mathsf{T}$ separately for the data sets $\mathcal{D}_{\text{full}}$ and $\mathcal{D}_{\text{frac}}$.

**Running Simulation:** The system is initialized at $\boldsymbol{x}_0$ from Table I which is generally not part of the training data. It uses estimates (11) to control the system based on (13). We run the simulation for $T_{\text{sim}} = 50s$ three times: Without any data $\mathcal{D}_{\text{none}} = \emptyset$, with $\mathcal{D}_{\text{frac}}$ and $\mathcal{D}_{\text{full}}$.

### B. Results

The estimates for $\hat{f}(\boldsymbol{x}), \hat{g}(\boldsymbol{x})$ along with the true function values $f\left(\boldsymbol{x}^{(i)}\right)$, $g\left(\boldsymbol{x}^{(i)}\right)$, $i = 1, \ldots, N_{\text{full}}$ are shown in Figs. 2 and 3. Note, that these points are not available for training, since only a noisy version of the sum $f\left(\boldsymbol{x}^{(i)}\right) + g\left(\boldsymbol{x}^{(i)}\right) u_0\left(\boldsymbol{x}^{(i)}\right)$ is observed. However, the identification performs well and distinguishes between the contribution of $\hat{f}(\boldsymbol{x})$ and $\hat{g}(\boldsymbol{x})$ to the measurements.

The plot in Fig. 4 shows the trajectories for the closed-loop system starting at $\boldsymbol{x}_0$ for the different training data
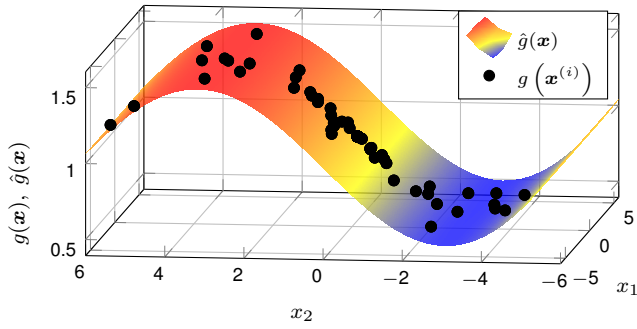
Fig. 3. The estimate $\hat{g}(\boldsymbol{x})$ (surface) along with the test points $g\left(\boldsymbol{x}^{(i)}\right)$, for $\mathcal{D}_{\text{full}}$ (black marks). Consider that $g\left(\boldsymbol{x}^{(i)}\right)$ are not available as measurements for identification but are only added here as reference.
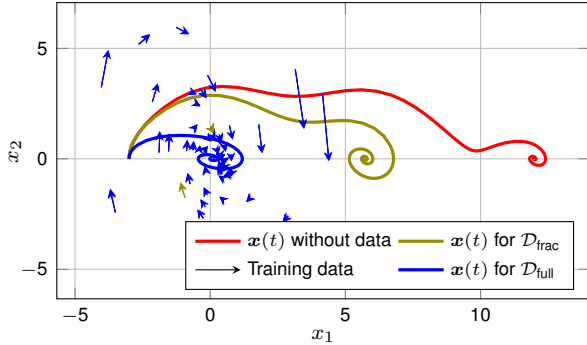


Fig. 4. Training data (arrows) and trajectories of system (19) with GP feedback linearization for no training data, a fraction of the training data $\mathcal{D}_{\text{frac}}$ and the full data set $\mathcal{D}_{\text{full}}$.

sets $\mathcal{D}_{\text{full}}, \mathcal{D}_{\text{frac}}, \mathcal{D}_{\text{none}}$. All trajectories are bounded and the ultimate bound contracts as more training data is added.

### C. Discussion

Gaussian processes turn out to be suitable since all prior assumptions are properly transfered into the model (Lemma 1) and the identification runs in closed-loop. Since no parametric structure is assumed for $f, g$, the model is very flexible and the data-driven framework increases its complexity as more data is available, which is the most significant advantage over classical system identification. Considering the difficulty mentioned in Remark 1, the structure in (19), ($g(\boldsymbol{x})$ only depends on $x_2$) is surely beneficial (simplifying the hyperparameter optimization) for the good result. However, this is common in many real-world systems.

The accuracy of the model depends on the distribution of the training data, which is determined by the initial controller $u_0$. There surely exist better methods for the exploration-exploitation trade-off, than the used PD controller, see [16], yet this is out of the scope of this work.

For the control side, we take advantage of an additional feature of the GP: The variance function allows high probability statements regarding the precision of the estimate without knowing the true function or an upper bound (only its RKHS norm). This is a significant advantage over neural networks or other robust control approaches. Knowing the precision allows to formulate the ultimate bound, to which

the system converges with high probability. Given a performance based control tasks, e.g. a convergence set size, the derived relationship allows to conclude whether the provided training data is sufficient or more information is needed.

## VI. CONCLUSION

This paper introduces a feedback linearization control law for initially unknown systems. We exploit a particular kernel structure of Gaussian processes to identify the control affine systems in a closed-loop setting. We show, that (under the specified assumption) the system is globally uniformly ultimately bounded. The advantage of GPs, the provided measure for the model fidelity, is exploited by deriving tighter bounds for the set of convergence under a specified probability. We show, that the control outcome improves as more training data is collected and validate this in simulation.

## REFERENCES

[1] H. K. Khalil and J. Grizzle, *Nonlinear systems*. Prentice hall New Jersey, 1996, vol. 3.
[2] L. Ljung, *System Identification*. NJ, USA: Prentice Hall PTR, 1998.
[3] G. Pillonetto, F. Dinuzzo, T. Chen, G. De Nicolao, and L. Ljung, "Kernel methods in system identification, machine learning and function estimation: A survey," *Automatica*, vol. 50, no. 3, pp. 657–682, 2014.
[4] J. M. Wang, D. J. Fleet, and A. Hertzmann, "Gaussian process dynamical models," in *Advances in Neural Information Processing Systems (NIPS)*, 2005, pp. 1441–1448.
[5] J. Kocijan, A. Girard, B. Banko, and R. Murray-Smith, "Dynamic systems identification with Gaussian processes," *Mathematical and Computer Modelling of Dynamical Systems*, vol. 11(4), 411-424, 2005.
[6] T. Beckers and S. Hirche, "Stability of Gaussian process state space models," in *European Control Conference*, 2016, pp. 2275–2281.
[7] ——, "Equilibrium distributions and stability analysis of Gaussian process state space models," in *Conference on Decision and Control (CDC)*, 2016, pp. 6355–6361.
[8] F. Berkenkamp and A. P. Schoellig, "Safe and robust learning control with Gaussian processes," in *European Control Conference (ECC)*, Jul. 2015, pp. 2496–2501.
[9] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Information-theoretic regret bounds for Gaussian process optimization in the bandit setting," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 3250–3265, May 2012.
[10] J.-J. E. Slotine and J. Karl Hedrick, "Robust input-output feedback linearization," *International Journal of control*, vol. 57, no. 5, pp. 1133–1139, 1993.
[11] A. L. D. Franco, H. Bourles, E. R. D. Pieri, and H. Guillard, "Robust nonlinear control associating robust feedback linearization and H-infinity control," *IEEE Transactions on Automatic Control*, vol. 51, no. 7, pp. 1200–1207, July 2006.
[12] A. Yesildirak and F. L. Lewis, "Feedback linearization using neural networks," *Automatica*, vol. 31, no. 11, pp. 1659–1664, 1995.
[13] G. Chowdhary, H. A. Kingravi, J. P. How, and P. A. Vela, "Bayesian nonparametric adaptive control using Gaussian processes," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 3, pp. 537–550, Mar. 2015.
[14] C. E. Rasmussen and C. K. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, Jan. 2006.
[15] D. Duvenaud, "Automatic model construction with Gaussian processes," Ph.D. dissertation, Computational and Biological Learning Laboratory, University of Cambridge, 2014.
[16] F. Berkenkamp, R. Moriconi, A. P. Schoellig, and A. Krause, "Safe learning of regions of attraction for uncertain, nonlinear systems with Gaussian processes," *arXiv preprint arXiv:1603.04915*, 2016.
[17] J. Umlauft, T. Beckers, M. Kimmel, and S. Hirche, "Feedback linearization using Gaussian processes," in *Conference on Decision and Control (CDC)*. IEEE, Dec 2017, pp. 5249–5255.