Stable Model-based Control with Gaussian Process Regression for Robot Manipulators

Thomas Beckers * **Jonas Umlauft** * **Sandra Hirche** *

* *Chair of Information-oriented Control (ITR), Department of Electrical and Computer Engineering, Technical University of Munich, Germany (e-mail: {t.beckers,jonas.umlauft,hirche}@tum.de)*

Abstract: Computed-torque control requires a very precise dynamical model of the robot for compensating the manipulator dynamics. This allows reduction of the controller's feedback gains resulting in disturbance attenuation and other advantages. Finding precise models for manipulators is often difficult with parametric approaches, e.g. in the presence of complex friction or flexible links. Therefore, we propose a novel computed-torque control law which consists of a PD feedback and a dynamic feed forward compensation part with Gaussian Processes. For this purpose, the nonparametric Gaussian Process regression infers the difference between an estimated and the true dynamics. In contrast to other approaches, we can guarantee that the tracking error is stochastically bounded. Furthermore, if the number of training points tends to infinity, the tracking error is asymptotically stable in the large. In simulation and with an experiment, we demonstrate the applicability of the proposed control law and that it outperforms classical computed-torque approaches in terms of tracking precision.

Keywords: Stochastic control, Stability of nonlinear systems, Data-based control Nonparametric methods, Adaptive system and control, robotic manipulators

1. INTRODUCTION

In the last decades, various robot control schemes have been proposed and most of them can be considered as a subset of computed-torque control laws. Computed-torque, a special case of feedback linearization, transforms the nonlinear system into an equivalent linear system through a change of variables and a suitable control input. Computed-torque control is able to derive very effective robot controllers that appear in robust, adaptive and learning control schemes [Siciliano et al., 2010]. With an exact model of the manipulator, this control law can compensate the robot dynamics to achieve a low gain feedback term which is beneficial in many ways: it allows safe physical human-robot interaction, reduces energy consumption, avoids disturbance attenuation in presence of noise and avoids the saturation of the actuators [Nguyen-Tuong and Peters, 2008]. Since the accuracy of the compensation depends on the precision of the model, the model building process is the key to achieve good performance.

The uncertainties of a model can be separated in structural and parametric variations. The structural uncertainties come from the lack of knowledge of the underlying true physics. The parametric uncertainties exist since the exact values of length, masses, etc. are often unknown. The classical approach is to derive a dynamic model from first order physic laws, e.g., with a CAD model of the manipulator and increase the feedback part of the control law to compensate for structural and parametric uncertainties until a desired performance is achieved [Spong and Vidyasagar, 2008]. But the increased gains are undesirable (as explained above) and therefore deriving more accurate models is of high importance.

Classical system identification algorithms for computed torque take advantage of the fact that the inertia parameters are linear with respect to the inverse dynamics [Slotine and Li, 1987]. Parametric models which cover all dynamics are hard to obtain, especially in the presents of friction, flexible links or environment interaction. Therefore, nonparametric learning approaches provide promising results [Deisenroth et al., 2015]. Gaussian Process regression (GPR) is a supervised learning technique with several advantages: it requires only a minimum of prior knowledge for arbitrary complex function, generalizes well even for little training data and has a precise trade-off between data fitting and smoothing [Rasmussen and Williams, 2006]. A Gaussian Process (GP) connects every point of a continuous input space with a normally distributed random variable. Any finite group of those infinitely many random variables follows a multivariate Gaussian distribution. Based on this, the result is a powerful tool for nonlinear function regression without the need of much prior knowledge. In contrast to many other regression techniques, e.g. neural networks, GP modeling provides not only a mean function but also a measure for the model fidelity based on the distance to training data. The output is a Gaussian distributed variable which is fully described by its mean and variance.

⋆ The research leading to these results has received funding from the European Research Council under the European Union Seventh Framework Program (FP7/2007-2013) / ERC Starting Grant "Control based on Human Models (con-humo)" agreement n°337654.

Alberto et al. [2014] present a computed-torque controller with Gaussian Process regression where the stiffness of the system depends on the variance of the learned model. The inherent learning of variable loads of the manipulator is done by Williams et al. [2009]. Nguyen-Tuong and Peters [2010] present a hybrid learning approach which incorporates model knowledge. The mentioned works show empirically promising results for GP-based control of robotic manipulators. However, they neither guarantee stability of the closed loop nor do they examine the influence of the training points mathematically.

The contribution of this paper is a novel computed torque control law for robotic manipulators using Gaussian Process regression, which guarantees stability and consistency of the regression. For this purpose, a GP learns the difference between an estimated model and the true robot from training trajectories. Afterwards, the control law uses GPR to compensate the unknown robot dynamics. The proposed method also abstains from feeding back joint accelerations (in comparison to, e.g., Alberto et al. [2014]) as these are difficult to measure directly and often inject noise. The derived method guarantees that the tracking error is stochastically bounded around zero independent of the number of training data. If the number of training points tends to infinity, this bound becomes arbitrary small and the tracking error is asymptotically stable in the large.

The paper is structured as follows: Section 2 introduces the model of a robotic manipulator and GPR. Section 3 proposes the control law, proofs it stability and explains training of GPs. An empirical evaluation based on simulation and experiment follows in Section 4.

2. BACKGROUND

2.1 Gaussian Process Regression

Let [1] $(\Omega, \mathcal{F}, P)$ be a probability space with the sample space Ω, the corresponding σ-algebra $\mathcal{F}$ and the probability measure P. The set $\mathcal{X} \subseteq \mathbb{R}^d$ with $d \in \mathbb{N}^*$ denotes the index set. A Gaussian Process is a discrete or real valued function $f_{GP}(\boldsymbol{x}, \omega)$ which is a measurable function of $\omega \in \Omega$ with $\boldsymbol{x} \in \mathcal{X}$. The process is fully described by a mean function $m(\boldsymbol{x})$ and a covariance function $k(\boldsymbol{x}, \boldsymbol{x}')$ since it is Gaussian distributed for any fixed $\boldsymbol{x}$. The GP is denoted by

$$f_{GP}(\boldsymbol{x}) \sim \mathcal{GP}(m(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}')), \qquad \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X},$$
$$m(\boldsymbol{x})\colon \mathcal{X} \to \mathbb{R},\ k(\boldsymbol{x}, \boldsymbol{x}')\colon \mathcal{X} \times \mathcal{X} \to \mathbb{R}.$$

The mean function is usually set to zero as no prior knowledge about the function is given. The covariance function is a measure for the interference of two states $(\boldsymbol{x}, \boldsymbol{x}')$. Probably the most widely used covariance function in GP modeling is the squared exponential covariance function with the set of hyperparameters $\varphi = \{\lambda \in \mathbb{R}_+^*, \sigma_f \in \mathbb{R}_+\}$ [Rasmussen and Williams, 2006]:

$$k_\varphi(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2 \exp\left(-\frac{\|\boldsymbol{x} - \boldsymbol{x}'\|^2}{2\lambda^2}\right)$$

The length-scale λ determines the number of expected upcrossing of the level zero in a unit interval by a zero-mean GP. The signal variance σ_f^2 describes the average distance of the function $f_{GP}(\boldsymbol{x})$ away from its mean. With this kernel any realization of $f_{GP}(\boldsymbol{x})$ is a smooth function, which makes it a suitable candidate for modeling physical dynamics.

In this paper, we use GPs for multivariate regression. Since the output of a GP is one dimensional, a regression over n outputs requires n GPs. Therefore, the vector valued function $\boldsymbol{m}(\cdot) = [m_1(\cdot), \ldots, m_n(\cdot)]^\top$ describes the mean functions for each component of a vector-valued $\boldsymbol{f}_{GP}$. The covariance functions for each state are bundled in the function $\boldsymbol{k}(\cdot,\cdot) = [k_{\varphi_1}(\cdot,\cdot), \ldots, k_{\varphi_n}(\cdot,\cdot)]^\top$ with

$$f_{GP,i}(\boldsymbol{x}) \sim \mathcal{GP}(m_i(\boldsymbol{x}), k_{\varphi_i}(\boldsymbol{x}, \boldsymbol{x}'))$$

for $i = 1, \ldots, n$ and the corresponding set of hyperparameters φ_i. The GP has to be provided with input/output pairs. For this purpose, we arrange the m training inputs $\{\boldsymbol{x}_i\}_{i=1}^m$ and outputs $\{\boldsymbol{y}_i\}_{i=1}^m$ pairs in an input training matrix $X = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m]$ and an output training matrix $Y = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_m]^\top$. Therefore, the training data set for the i-th GP is described by $\mathcal{D}_i = \{X, Y_{:,i}\}$ where $Y_{:,i}$ is the i-th column of the matrix Y. The hyperparameters φ_i are trained through likelihood optimization, thus by maximizing the probability of the seen data to occur given the current parameters and input values

$$\varphi_i^* = \arg\max_{\varphi_i} \log P(Y_{:,i}|X, \varphi_i).$$

The predicted output $\boldsymbol{y}^* \in \mathbb{R}^n$ for a test value $\boldsymbol{x}^*$ is a Gaussian distributed variable. With the assumption that the mean functions of the GPs are set to zero, a prediction of the i-th component of $\boldsymbol{y}^*|\boldsymbol{x}^*$ is obtained with

$$\begin{aligned} y_i^* &\sim \mathcal{N}\left(\mu_i(\boldsymbol{y}^*), \mathrm{var}_i(\boldsymbol{y}^*)\right), \\ \mu_i(\boldsymbol{y}^*) &= \boldsymbol{k}_{\varphi_i}(\boldsymbol{x}^*, X)^\top (K_{\varphi_i}(X, X) + I_m \sigma_{ni}^2)^{-1} Y_{:,i}, \\ \mathrm{var}_i(\boldsymbol{y}^*) &= k_{\varphi_i}(\boldsymbol{x}^*, \boldsymbol{x}^*) - \boldsymbol{k}_{\varphi_i}(\boldsymbol{x}^*, X)^\top \\ &\quad (K_{\varphi_i}(X, X) + I_m \sigma_{ni}^2)^{-1} \boldsymbol{k}_{\varphi_i}(\boldsymbol{x}^*, X), \end{aligned} \tag{1}$$

where $\mu_i(\cdot)$ is the mean and $\mathrm{var}_i(\cdot)$ the variance of the random variable. The matrix $K_{\varphi_i}(X, X)$ denotes the concatenation of pairwise evaluation of all input data points and $\boldsymbol{k}_{\varphi_i}(\boldsymbol{x}, X)$ the vector-valued extended covariance function with the set of hyperparameters φ_i. The variable $\sigma_{n_i}^2 \in \mathbb{R}_+$ is the variance of the input data for all $i \in \{1, \ldots, m\}$. It increases the numerical stability of the matrix inversion. The n normally distributed components are combined in a multi-variable distribution

$$\begin{aligned} \boldsymbol{y}^* &\sim \mathcal{N}\left(\boldsymbol{\mu}(\boldsymbol{y}^*), \mathrm{var}(\boldsymbol{y}^*)\right), \\ \boldsymbol{\mu}(\boldsymbol{y}^*) &= [\mu_1(\boldsymbol{y}^*), \ldots, \mu_n(\boldsymbol{y}^*)]^\top, \\ \Sigma(\boldsymbol{y}^*) &= \mathrm{diag}(\mathrm{var}_1(\boldsymbol{y}^*), \ldots, \mathrm{var}_n(\boldsymbol{y}^*)). \end{aligned}$$

2.2 Dynamic Model of a Robot Manipulator

The dynamics of an n-link rigid manipulator can be written as

$$H(\boldsymbol{q})\ddot{\boldsymbol{q}} + C(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{\tau}, \tag{2}$$

where $\boldsymbol{q} \in \mathbb{R}^n$ is the vector of continuous joint displacements and $\boldsymbol{\tau} \in \mathbb{R}^n$ is the vector of applied joint

[1] **Notation**: Vectors and vector-valued functions are denoted with bold characters. Capital letters describe matrices. The expression $\mathcal{N}(\mu, \Sigma)$ denotes the normal distribution with mean μ and covariance Σ. The Euclidean norm is given by $\|\cdot\|$. The mean and variance of a random variable is written as $\mu(\cdot)$ or $\mathrm{E}(\cdot)$ and $\mathrm{var}(\cdot)$, respectively. The minimum singular value of a matrix is denoted with $\sigma_{min}(\cdot)$. I_n denotes the $n \times n$ identity matrix.

torques. The matrix $H(\boldsymbol{q}) \in \mathbb{R}^{n\times n}$ is the manipulator's inertia, $C(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}}$ in $\mathbb{R}^n$ is the vector of centripetal and Coriolis torques, and $\boldsymbol{g}(\boldsymbol{q}) \in \mathbb{R}^n$ includes gravity terms and other forces which act at the joints which have the following properties:

Property 1 (Structural properties)

- *$H(\boldsymbol{q})$ is symmetric and positive definite.*
- *$\dot{H}(\boldsymbol{q})-2C(\boldsymbol{q},\dot{\boldsymbol{q}}) \in \mathbb{R}^{n\times n}$ is a skew-symmetric matrix.*

Additionally, for robots with only rotational joints, the following properties hold [Ghorbel et al., 1993]:

Property 2 (Boundedness and Linearity)

- *The inertia matrix is bounded and Lipschitz continuous, i.e. $\|H(\boldsymbol{q})\| < \infty$ and $\|H(\boldsymbol{q})-H(\boldsymbol{q}')\| \leq L\|\boldsymbol{q}-\boldsymbol{q}'\|$ with $L > 0$, for all $\boldsymbol{q}, \boldsymbol{q}' \in \mathbb{R}^n$.*
- *The matrix $C(\boldsymbol{q},\dot{\boldsymbol{q}})$ is bounded in $\boldsymbol{q}$ and linear in $\dot{\boldsymbol{q}}$, i.e. $\|C(\boldsymbol{q},\dot{\boldsymbol{q}})\| \leq c_C\|\dot{\boldsymbol{q}}\|$ and $C(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{p}} = C(\boldsymbol{q},\dot{\boldsymbol{p}})\dot{\boldsymbol{q}}$ for all $\boldsymbol{q},\dot{\boldsymbol{q}},\dot{\boldsymbol{p}} \in \mathbb{R}^n$ and $c_C \in \mathbb{R}_+$.*

2.3 Stochastic Stability Definitions

Since we propose the use of GPs (a stochastic process), we quickly review the concept of stochastic stability and boundedness of the stochastic differential equation

$$d\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x},t)dt + G(\boldsymbol{x},t)d\boldsymbol{\omega} \quad (3)$$
$$\text{with } \boldsymbol{f}\colon \mathbb{R}^{n_1}\times\mathbb{R}_+ \to \mathbb{R}^{n_1},\ G\colon \mathbb{R}^{n_1}\times\mathbb{R}_+ \to \mathbb{R}^{n_1\times n_2},$$

where $\boldsymbol{\omega}$ indicates the Brownian motion and $n_1, n_2 \in \mathbb{N}$.

Definition 1 (Stochastic Sample Path Boundedness)
Let there exist a ball $B = \{\|\boldsymbol{x}\| \leq r | \boldsymbol{x} \in \mathbb{R}^{n_1}, r > 0\}$ and a time τ_1 denoting the first exit time from $\mathbb{R}^{n_1}\backslash B$ for the solution $\boldsymbol{x}(t)$ where $\boldsymbol{x_0} \in \mathbb{R}^{n_1}\backslash B$. The system (3) is stochastically sample path bounded, if for each $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$P\left(\sup_{0\leq t\leq\tau_1} \|\boldsymbol{x}(t)\| \leq \delta\right) > 1-\varepsilon.$$

According to Gard [1988] this is shown as follows:

Theorem 1. Let there exist a proper Lyapunov function $V \in C^2$ for which the drift operator

$$LV(\boldsymbol{x}) = \frac{\partial V}{\partial \boldsymbol{x}}\boldsymbol{f} + \frac{1}{2}\operatorname{Tr}\left(G^\top \frac{\partial^2 V}{\partial \boldsymbol{x}\partial \boldsymbol{x}} G\right) \leq 0$$

holds for $\boldsymbol{x} \in \mathbb{R}^{n_1}\backslash B$. Than the solution of (3) is stochastic sample path bounded with $B = \{\|\boldsymbol{x}\| \leq r | \boldsymbol{x} \in \mathbb{R}^{n_1}, r > 0\}$.

A stronger stability criteria is given as follows:

Definition 2 (Stochastic Asymptotic Stability in the large)
The system (3) is stochastically asymptotically stable in the large if it is stochastically stable in probability and for all $\boldsymbol{x}_0 \in \mathbb{R}^{n_1}$ holds

$$P\left(\lim_{t\to\infty} \boldsymbol{x}(t) = \boldsymbol{0}\right) = 1.$$

Stability in the large is shown as following [Mao, 2007]:

Theorem 2. If there exists a positive-definite radially unbounded function $V(\boldsymbol{x}) \in \mathcal{C}^2$ such that the drift operator LV is negative-definite, then the trivial solution of (3) is stochastically asymptotically stable in the large.

We will use these theorems to show that the tracking error is stochastically sample path bounded and for an infinite number of training points the tracking error is asymptotic stable in the large. Since stability in the large considers all realization of the stochastic process, all conclusions also hold if the mean function of the control law is applied in a deterministic setting.

3. GAUSSIAN PROCESSES FOR MODIFIED COMPUTED-TORQUE

The augmented PD control law from Siciliano et al. [2010] requires knowledge about the model, i.e the matrices H, C and $\boldsymbol{g}$. To avoid this requirement, we use a modified version of the control law which learns the difference between the model, given by the estimates $\hat{H}, \hat{C}$ and $\hat{\boldsymbol{g}}$ and the true robot dynamics. The estimates $\hat{H}, \hat{C}, \hat{\boldsymbol{g}}$ can be obtained from CAD data or from standard identification procedures.

3.1 Conditions

The goal is to design a control law $\boldsymbol{u}_c$ which tracks the desired joint trajectory $\boldsymbol{q}_d, \dot{\boldsymbol{q}}_d, \ddot{\boldsymbol{q}}_d$ under the following conditions:

- C1 The desired trajectory satisfies $\|\boldsymbol{q_d}\| < c_q,\ \|\dot{\boldsymbol{q}}_{\boldsymbol{d}}\| < c_{\dot{q}}$, and $\|\ddot{\boldsymbol{q}}_{\boldsymbol{d}}\| < c_{\ddot{q}}$ with $c_q, c_{\dot{q}}, c_{\ddot{q}} \in \mathbb{R}_+$.
- C2 The feedback gain matrices K_d and K_p are positive definite and the smallest singular value of K_d is larger than $\beta \in \mathbb{R}_+$, thus $\sigma_{min}(K_d) > \beta$.
- C3 The norm of the model error is affinely bounded by the norm of the angular velocity, i.e. $\|H(\boldsymbol{q})\ddot{\boldsymbol{q}}_{\boldsymbol{d}} + C(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}}_{\boldsymbol{d}} + \boldsymbol{g}(\boldsymbol{q}) - \hat{H}(\boldsymbol{q})\ddot{\boldsymbol{q}}_{\boldsymbol{d}} - \hat{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}}_{\boldsymbol{d}} - \hat{\boldsymbol{g}}(\boldsymbol{q})\| \leq \alpha + \beta\|\dot{\boldsymbol{q}}\|$ for all $\boldsymbol{q}, \dot{\boldsymbol{q}} \in \mathbb{R}^n$ with $\alpha, \beta \in \mathbb{R}_+$, and continuous regarding to $\ddot{\boldsymbol{q}}_{\boldsymbol{d}}, \dot{\boldsymbol{q}}_{\boldsymbol{d}}, \dot{\boldsymbol{q}}, \boldsymbol{q}$.

The conditions in C1, i.e. bounded reference motion trajectories, are a very natural assumption and do not pose any restriction in practice. Also C2 does not restrict the applicability of our approach but must be kept in mind during the design of the controller. From a practical point of view, C3 states that the dynamics which are not modeled by (2) can at most depend linearly on the joint velocity. If there is a known range of uncertainty only in the inertia parameter, the values α and β can be computed using the approach of Takegaki and Arimoto [1981]. Since the payload is the major reason for the uncertainty, this approach is suitable for most application scenarios.

3.2 Control Law

The following theorem proposes a control law to ensure a bounded tracking error under the proposed conditions.

Theorem 3. Assume an n-link rigid manipulator with only rotational joints

$$\boldsymbol{\tau} = H(\boldsymbol{q})\ddot{\boldsymbol{q}} + C(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \boldsymbol{g}(\boldsymbol{q}), \quad (4)$$

for which Properties 1 and 2 and C1-C3 hold. Given an estimated model of the manipulator

$$\hat{\boldsymbol{\tau}} = \hat{H}(\boldsymbol{q})\ddot{\boldsymbol{q}} + \hat{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \hat{\boldsymbol{g}}(\boldsymbol{q})$$

with control law

$$\boldsymbol{u_c} = \hat{H}(\boldsymbol{q})\ddot{\boldsymbol{q}}_{\boldsymbol{d}} + \hat{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}}_{\boldsymbol{d}} + \hat{\boldsymbol{g}}(\boldsymbol{q}) + \boldsymbol{f}_{GP}(\boldsymbol{q_c}) - K_d\dot{\boldsymbol{e}} - K_p\boldsymbol{e}, \quad (5)$$

shown in Fig. 1, where $\boldsymbol{q_c} = [\ddot{\boldsymbol{q}}_{\boldsymbol{d}}^\top, \dot{\boldsymbol{q}}_{\boldsymbol{d}}^\top, \boldsymbol{q}^\top]^\top$ and $\boldsymbol{f}_{GP}$ contains n posteriors of GPs with squared exponential kernel

$$\boldsymbol{f}_{GP}(\boldsymbol{q_c}) = \begin{bmatrix} f_{GP,1} \sim \mathcal{N}(\mu_1(\boldsymbol{q_c}), \mathrm{var}_1(\boldsymbol{q_c})) \\ \vdots \\ f_{GP,n} \sim \mathcal{N}(\mu_n(\boldsymbol{q_c}), \mathrm{var}_n(\boldsymbol{q_c})) \end{bmatrix}.$$

Then, the tracking error $\boldsymbol{e} = \boldsymbol{q} - \boldsymbol{q_d}$ is stochastically sample path bounded.

The stability proof of the control law proposed in Theorem 3 bases on the work in Whitcomb et al. [1993].

Proof. With $\boldsymbol{\tau} = \boldsymbol{u_c}$ the closed loop system is given by

$$\begin{aligned}\ddot{\boldsymbol{e}} = H(\boldsymbol{q})^{-1}(\hat{H}(\boldsymbol{q})\ddot{\boldsymbol{q}}_{\boldsymbol{d}} + \hat{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}}_{\boldsymbol{d}} - C(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \hat{\boldsymbol{g}}(\boldsymbol{q}) \\ - \boldsymbol{g}(\boldsymbol{q}) + \boldsymbol{f}_{GP}(\boldsymbol{q_c}) - K_d\dot{\boldsymbol{e}} - K_p\boldsymbol{e}) - \ddot{\boldsymbol{q}}_{\boldsymbol{d}},\end{aligned}$$

since the matrix $H(\boldsymbol{q})$ is always non-singular. The posterior of the GP $\boldsymbol{f}_{GP}(\boldsymbol{q_c})$ can be split in a drift $\boldsymbol{\mu}(\boldsymbol{q_c})$ and a diffusion term $\Sigma(\boldsymbol{q_c})\boldsymbol{w}$

$$\begin{aligned}\boldsymbol{f}_{GP}(\boldsymbol{q_c}) &= \boldsymbol{\mu}(\boldsymbol{q_c}) + \Sigma(\boldsymbol{q_c})\boldsymbol{w}, \\ \boldsymbol{\mu}(\boldsymbol{q_c}) &= \begin{bmatrix} \boldsymbol{k}_{\varphi_1}^\top (K_{\varphi_1} + I_m\sigma_{n1}^2)^{-1} Y_{:,1} \\ \vdots \\ \boldsymbol{k}_{\varphi_n}^\top (K_{\varphi_n} + I_m\sigma_{nn}^2)^{-1} Y_{:,n} \end{bmatrix}, \\ \Sigma(\boldsymbol{q_c}) &= \mathrm{diag} \begin{bmatrix} k_{\varphi_1} - \boldsymbol{k}_{\varphi_1}^\top (K_{\varphi_1} + I_m\sigma_{n1}^2)^{-1}\boldsymbol{k}_{\varphi_1} \\ \vdots \\ k_{\varphi_n} - \boldsymbol{k}_{\varphi_n}^\top (K_{\varphi_n} + I_m\sigma_{nn}^2)^{-1}\boldsymbol{k}_{\varphi_n} \end{bmatrix}^{\frac{1}{2}},\end{aligned}$$

with an n-dimensional standard Brownian noise vector $\boldsymbol{w} = [w_1 \dots w_n]^\top$. We reformulate the closed loop system with a drift and a diffusion term

$$\begin{aligned}\frac{d}{dt}\begin{bmatrix}\dot{\boldsymbol{e}} \\ \boldsymbol{e}\end{bmatrix} &= \begin{bmatrix} H^{-1}(\hat{H}\ddot{\boldsymbol{q}}_{\boldsymbol{d}} + \hat{C}\dot{\boldsymbol{q}}_{\boldsymbol{d}} - C\dot{\boldsymbol{q}} + \hat{\boldsymbol{g}} - \boldsymbol{g} \\ +\boldsymbol{\mu}(\boldsymbol{q_c}) - K_d\dot{\boldsymbol{e}} - K_p\boldsymbol{e}) - \ddot{\boldsymbol{q}}_{\boldsymbol{d}} \\ \dot{\boldsymbol{e}} \end{bmatrix} \\ &\quad + \begin{bmatrix} H^{-1}\Sigma(\boldsymbol{q_c}) \\ 0 \end{bmatrix}\boldsymbol{w} := \boldsymbol{f}(\boldsymbol{\xi}, t) + G(\boldsymbol{\xi})\boldsymbol{w}, \qquad (6)\end{aligned}$$

where the vector $\boldsymbol{\xi} = [\boldsymbol{q}^\top, \dot{\boldsymbol{q}}^\top]^\top \in \mathcal{X} \subset \mathbb{R}^{2n}$. For the stability analysis of this stochastic differential equation we use the differential generator L which maps C^2 functions $V\colon \mathcal{X} \to \mathbb{R}$ to C^0 functions $LV\colon \mathcal{X} \to \mathbb{R}$ given by Theorem 2. Assume the following Lyapunov function

$$V(\boldsymbol{\xi}) = \underbrace{\frac{1}{2}\dot{\boldsymbol{e}}^\top H\dot{\boldsymbol{e}} + \frac{1}{2}\boldsymbol{e}^\top K_p\boldsymbol{e}}_{V_0(\boldsymbol{\xi})} + \underbrace{\varepsilon\boldsymbol{e}^\top H\dot{\boldsymbol{e}}}_{V_{\mathrm{mix}}(\boldsymbol{\xi})}, \qquad (7)$$

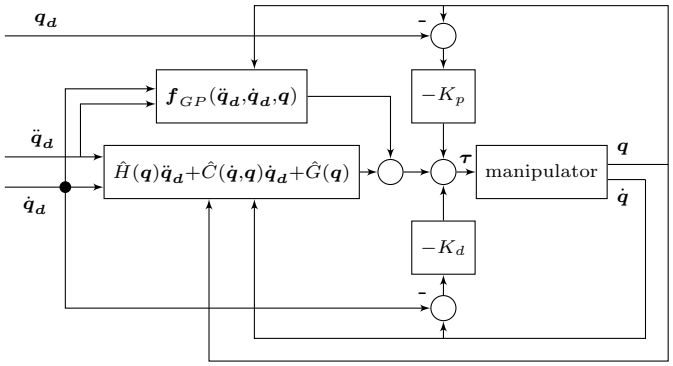

Fig. 1. Structure of the proposed closed loop controller (5).

which contains the two terms V_0 and V_{mix}. We start with the computation of LV_0 using the property that the matrix H^{-1} is symmetric

$$\begin{aligned}LV_0 = \dot{\boldsymbol{e}}^\top H\ddot{\boldsymbol{e}} + \frac{1}{2}\dot{\boldsymbol{e}}^\top\dot{H}\dot{\boldsymbol{e}} + \dot{\boldsymbol{e}}^\top K_p\boldsymbol{e} \\ + \frac{1}{2}\mathrm{Tr}\,\Sigma^\top(\boldsymbol{q_c})H^{-1}\Sigma(\boldsymbol{q_c}).\end{aligned}$$

Now, the acceleration error $\ddot{\boldsymbol{e}}$ of (6) is substituted.

$$\begin{aligned}LV_0 &= \dot{\boldsymbol{e}}^\top(\hat{H}\ddot{\boldsymbol{q}}_{\boldsymbol{d}} + \hat{C}\dot{\boldsymbol{q}}_{\boldsymbol{d}} - C\dot{\boldsymbol{q}} + \hat{\boldsymbol{g}} - \boldsymbol{g} - K_d\dot{\boldsymbol{e}} \\ &\quad - K_p\boldsymbol{e} - H\ddot{\boldsymbol{q}}_{\boldsymbol{d}} + \boldsymbol{\mu}(\boldsymbol{q_c})) + \dot{\boldsymbol{e}}^\top\left(\frac{1}{2}(\dot{H} - 2C) + C\right)\dot{\boldsymbol{e}} \\ &\quad + \dot{\boldsymbol{e}}^\top K_p\boldsymbol{e} + \frac{1}{2}\mathrm{Tr}\,\Sigma^\top(\boldsymbol{q_c})H^{-1}\Sigma(\boldsymbol{q_c}).\end{aligned}$$

Since Property 1 implies the skew-symmetry of $\dot{H} - 2C$, this term can be canceled out and LV_0 is given by

$$\begin{aligned}LV_0 = -\dot{\boldsymbol{e}}^\top K_d\dot{\boldsymbol{e}} + \dot{\boldsymbol{e}}^\top(\tilde{H}\ddot{\boldsymbol{q}}_{\boldsymbol{d}} + \tilde{C}\dot{\boldsymbol{q}}_{\boldsymbol{d}} + \tilde{\boldsymbol{g}} + \boldsymbol{\mu}(\boldsymbol{q_c})) \\ + \frac{1}{2}\mathrm{Tr}\,\Sigma^\top(\boldsymbol{q_c})H^{-1}\Sigma(\boldsymbol{q_c}), \qquad (8)\end{aligned}$$

with the difference between the manipulator model and the estimated model matrices defined by

$$\begin{aligned}\tilde{H}(\boldsymbol{q}) &= \hat{H}(\boldsymbol{q}) - H(\boldsymbol{q}), \\ \tilde{C}(\boldsymbol{q},\dot{\boldsymbol{q}}) &= \hat{C}(\boldsymbol{q},\dot{\boldsymbol{q}}) - C(\boldsymbol{q},\dot{\boldsymbol{q}}), \\ \tilde{\boldsymbol{g}}(\boldsymbol{q}) &= \hat{\boldsymbol{g}}(\boldsymbol{q}) - \boldsymbol{g}(\boldsymbol{q}).\end{aligned}$$

The first summand of LV_0 is negative definite but the definiteness of the second part is indeterminate. However, condition C3 guarantees that $\|\tilde{H}\ddot{\boldsymbol{q}}_{\boldsymbol{d}} + \tilde{C}\dot{\boldsymbol{q}}_{\boldsymbol{d}} + \tilde{\boldsymbol{g}}\|$ is bounded by an affine function $\alpha + \beta\|\dot{\boldsymbol{q}}\|$. The mean prediction $\|\boldsymbol{\mu}(\boldsymbol{q_c})\|$ and the corresponding variance $\|\Sigma(\boldsymbol{q_c})\|$ is also bounded [Beckers and Hirche, 2016]. Therefore, it is possible to find an upper bound for the second and third part of (8)

$$\|\dot{\boldsymbol{e}}^\top(\tilde{H}\ddot{\boldsymbol{q}}_{\boldsymbol{d}} + \tilde{C}\dot{\boldsymbol{q}}_{\boldsymbol{d}} + \tilde{\boldsymbol{g}} + \boldsymbol{\mu}(\boldsymbol{q_c}))\| \leq \|\dot{\boldsymbol{e}}\|(\alpha + c_\mu + \beta\|\dot{\boldsymbol{q}}\|). \qquad (9)$$

Since $\|H(\dot{\boldsymbol{q}})\|$ is bounded and the matrix is always non-singular, the inverse $\|H^{-1}(\dot{\boldsymbol{q}})\|$ is also bounded

$$\frac{1}{2}\mathrm{Tr}\,\Sigma^\top(\boldsymbol{q_c})H^{-1}\Sigma(\boldsymbol{q_c}) \leq c_g \in \mathbb{R}_+.$$

These results are used for the estimation of an upper bound for LV_0

$$LV_0 \leq -\dot{\boldsymbol{e}}^\top K_d\dot{\boldsymbol{e}} + \|\dot{\boldsymbol{e}}\|\beta\|\dot{\boldsymbol{e}}\| + \|\dot{\boldsymbol{e}}\|(\alpha + \beta c_{\dot{q}} + c_\mu) + c_g. \qquad (10)$$

If $\sigma_{min}(K_d) > \beta$, the first part dominates the second part of the equation and the sum is negative definite. Since it is negative definite, the quadratic part of LV_0, i.e. $-\dot{\boldsymbol{e}}^\top K_d\dot{\boldsymbol{e}} + \|\dot{\boldsymbol{e}}\|\beta\|\dot{\boldsymbol{e}}\|$, dominates the linear and the constant part for $\dot{\boldsymbol{e}} \to \infty$. Thus the following holds

$$\lim_{\|\dot{\boldsymbol{e}}\|\to\infty} LV_0(\boldsymbol{\xi}, t) = -\infty,$$

for all $\boldsymbol{e} \in \mathbb{R}^n$. With the continuity of V_0, there exist a ball $B_{\dot{\boldsymbol{e}}} = \{\|\dot{\boldsymbol{e}}\| \leq \delta_{\dot{e}}\}$ with the property that $LV_0(\boldsymbol{\xi}, t) < 0$ if $\dot{\boldsymbol{e}} \in \mathbb{R}^n \setminus B_{\dot{\boldsymbol{e}}}$. In other words, LV_0 is negative outside $B_{\dot{\boldsymbol{e}}}$ and therefore $\dot{\boldsymbol{e}}$ is stochastically sample path bounded. Boundedness of the tracking error $\boldsymbol{e}$ can not be guaranteed so far because LV_0 does not depend on $\boldsymbol{e}$. Therefore, the second part V_{mix} of Lyapunov function (7) is included

$$\begin{aligned}LV_{\mathrm{mix}} &= \varepsilon(\dot{\boldsymbol{e}}^\top H\dot{\boldsymbol{e}} + \boldsymbol{e}^\top\dot{H}\dot{\boldsymbol{e}} + \boldsymbol{e}^\top H\ddot{\boldsymbol{e}}) \\ &= \varepsilon\dot{\boldsymbol{e}}^\top H\dot{\boldsymbol{e}} + \varepsilon\boldsymbol{e}^\top\dot{H}\dot{\boldsymbol{e}} + \varepsilon\boldsymbol{e}^\top(-K_p\boldsymbol{e} - K_d\dot{\boldsymbol{e}} \\ &\quad + \tilde{H}\ddot{\boldsymbol{q}}_{\boldsymbol{d}} + \tilde{C}\dot{\boldsymbol{q}}_{\boldsymbol{d}} - C\dot{\boldsymbol{e}} + \tilde{\boldsymbol{g}} + \boldsymbol{\mu}(\boldsymbol{q_c})).\end{aligned}$$

After rewriting the equation

$$LV_{\text{mix}} = \varepsilon\Big(\underbrace{-\boldsymbol{e}^\top K_p \boldsymbol{e}}_{LV_1} + \underbrace{\dot{\boldsymbol{e}}^\top H \dot{\boldsymbol{e}}}_{LV_2} + \underbrace{\boldsymbol{e}^\top (\dot{H} - C - K_d)\dot{\boldsymbol{e}}}_{LV_3} + \underbrace{\boldsymbol{e}^\top \left(\tilde{H}\ddot{\boldsymbol{q}}_{\boldsymbol{d}} + \tilde{C}\dot{\boldsymbol{q}}_{\boldsymbol{d}} + \tilde{\boldsymbol{g}} + \boldsymbol{\mu}(\boldsymbol{q_c}) \right)}_{LV_4} \Big),$$

the term is analyzed: The first summand LV_1 is negative definite from C2. The second part LV_2 sums up with $-\dot{\boldsymbol{e}}^\top K_d \dot{\boldsymbol{e}}$ of LV_0, see (10), but since ε can be arbitrary small the sum is still negative definite. For the cross term LV_3 it is sufficient to show the boundedness of the operator norm of $(\dot{H} - C - K_d)$. The upper bound

$$LV_3 \leq \left(\|K_d\| + \frac{5}{2} \left\| \frac{\partial H}{\partial \boldsymbol{q}} \right\| \|\dot{\boldsymbol{q}}\| \right) \|\boldsymbol{e}\| \|\dot{\boldsymbol{e}}\| \leq c_{V3} \|\boldsymbol{e}\| \|\dot{\boldsymbol{e}}\|,$$

with $c_{V3} \in \mathbb{R}_+$ can be shown by using the chain rule and the dependency between $\dot{H}$ and C. The partial derivation $\left\| \frac{\delta H}{\delta \boldsymbol{q}} \right\|$ is a bounded operator since $H(\boldsymbol{q})$ is Lipschitz continuous. So LV_3 is a ε-size bounded operator on $\boldsymbol{e}$ and $\dot{\boldsymbol{e}}$ which preserve the negative definiteness. The last part LV_4 is bounded by

$$LV_4 \leq \|\boldsymbol{e}\| \beta \|\dot{\boldsymbol{e}}\| + \|\boldsymbol{e}\| \underbrace{(\alpha + \beta c_{\dot{q}+} c_\mu)}_{c_{V4}}$$

with $c_{V4} \in \mathbb{R}_+$ which is analog to (9). After combining the parts, the upper bound for the diffusion operator L of the overall Lyapunov function V is given by:

$$\begin{aligned} LV &= -\dot{\boldsymbol{e}}^\top (K_d - \varepsilon H)\dot{\boldsymbol{e}} - \varepsilon \boldsymbol{e}^\top K_p \boldsymbol{e} + \varepsilon \boldsymbol{e}^\top (\dot{H} - C - K_d)\dot{\boldsymbol{e}} \\ &\quad + (\dot{\boldsymbol{e}}^\top + \varepsilon \boldsymbol{e}^\top)(\tilde{H}\ddot{\boldsymbol{q}}_{\boldsymbol{d}} + \tilde{C}\dot{\boldsymbol{q}}_{\boldsymbol{d}} + \tilde{\boldsymbol{g}} + \boldsymbol{\mu}(\boldsymbol{q_c})) \\ &\quad + \frac{1}{2} \operatorname{Tr} \Sigma^\top(\boldsymbol{q_c}) H^{-1} \Sigma(\boldsymbol{q_c}) \qquad (11) \\ &\leq -\dot{\boldsymbol{e}}^\top (K_d - \varepsilon H)\dot{\boldsymbol{e}} + \|\dot{\boldsymbol{e}}\| \beta \|\dot{\boldsymbol{e}}\| + (c_{V4} + c_g)\|\dot{\boldsymbol{e}}\| \\ &\quad - \varepsilon \boldsymbol{e}^\top K_p \boldsymbol{e} + \varepsilon (c_{V3} + \beta)\|\boldsymbol{e}\| \|\dot{\boldsymbol{e}}\| + \varepsilon c_{V4} \|\boldsymbol{e}\|. \end{aligned}$$

In comparison to (10), this term also includes a dominant quadratic part which depends on the error $\boldsymbol{e}$. Therefore, it is possible to find an $\varepsilon > 0$ which creates the ball

$$B_{\boldsymbol{\xi}} = \{\|\boldsymbol{\xi}\| \leq \delta_{\dot{e}e}\}$$
$$\text{with } LV < 0 \text{ for } \boldsymbol{\xi} \in \mathbb{R}^{2n} \setminus B_{\boldsymbol{\xi}}.$$

Consequently, the tracking error is stochastically sample path bounded and enters in a finite time the set $B_{\boldsymbol{\xi}}$. ■

Remark 1
The proposed control law uses Gaussian Processes with zero mean functions which is reasonable if no prior knowledge about the model error is given. However, if a priori knowledge is available, the Gaussian Process regression can be supported by a nonzero mean function. A bounded tracking error is preserved if (9) remains bounded which is fulfilled as long as the mean function is bounded [Beckers and Hirche, 2016].

The proof applies Theorem 2 to show that the tracking error is stochastically sample path bounded with the ball $B_{\boldsymbol{\xi}}$. For a radially unbounded, positive-definite Lyapunov function, it shows that the drift operator is negative outside of this ball and therefore the tracking error enters the ball in a finite time.
Equation (8) shows the need of C3 to ensures the global negative definiteness of LV_0 outside the ball. However, with less restrictions on the model error, it is possible to find local areas of boundedness which requires appropriate initial states. Important to note here is the stochastic nature of the control low $\boldsymbol{u_c}$ as it uses with $\boldsymbol{f}_{GP}$ a stochastic process. Nevertheless, its deterministic counterpart (the mean function) also results in the desired property:

Corollary 4. The rigid manipulator described by (4) with control law (5) results in a bounded tracking error if the stochastic process $\boldsymbol{f}_{GP}(\boldsymbol{q_c})$ is replaced by its deterministic posterior mean function $\boldsymbol{\mu}(\boldsymbol{q_c})$ as defined in (1).

This corollary directly follows from Theorem 3 since with stochastic sample path boundedness it must hold for all realizations of the stochastic control law. The mean function is simply one of these realizations.

3.3 Training

While the previous section proposed the control law which guarantees stochastic sample path boundedness around the origin, this section shows how this can tend to asymptotic stability through training of the GP. So practically speaking, training aims to shrink the radius of the ball $B_{\boldsymbol{\xi}}$ as the number of training points increases. For this purpose, the Gaussian Process learns the difference between the real and the estimated dynamics of the manipulator

$$\tilde{\boldsymbol{\tau}} = \boldsymbol{\tau} - \hat{\boldsymbol{\tau}} = -\tilde{H}(\boldsymbol{q})\ddot{\boldsymbol{q}} - \tilde{C}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}} - \tilde{\boldsymbol{g}}(\boldsymbol{q}). \qquad (12)$$

Figure 2 shows how to generate training pairs of various joint states $\{\ddot{\boldsymbol{q}}_i, \dot{\boldsymbol{q}}_i, \boldsymbol{q}_i\}_{i=1}^m$ and differences between the applied torque and the estimated torque of the model $\{\tilde{\boldsymbol{\tau}}_i\}_{i=1}^m$. The manipulator can be excited directly with a torque or with a stabilizing controller to drive the manipulator in the desired states. It is advisable that the area of training points is similar to the desired operation area of the manipulator. An appropriate choice of training points inside this area can be done, for example, with the Bayesian optimization method where the next training point is set to the position of maximum variance, as proposed in Sui et al. [2015]. After recording the training pairs, the Gaussian Process can be trained using likelihood optimization.
To achieve a asymptotically stable tracking error, the mean function $\boldsymbol{\mu}(\boldsymbol{q_c})$ of the GP must cancel out the model error $\tilde{\boldsymbol{\tau}}$ given in (12). We introduce the following lemma for the analysis of the convergence.

Lemma 1 (Consistency)
Let $\boldsymbol{f}_0 : \mathbb{R}^n \to \mathbb{R}^n$ a continuous function. A set of m training points is distributed uniformly in a bounded space or in a stochastic way with a nonzero density function.
A Gaussian Process $\boldsymbol{f}_{GP}$ with squared exponential kernel is consistent, i.e.

$$\mathrm{E}_{f_0} \left\| \mathrm{E}\left(\boldsymbol{f}_{GP}\right) - \boldsymbol{f}_0 \right\|^2 \overset{m \to \infty}{\longrightarrow} 0,$$

on the compact domain [Vaart and Zanten, 2011].

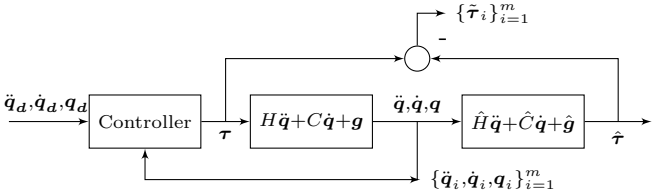

Fig. 2. The structure for generating the training data set $\mathcal{D} = \{[\ddot{\boldsymbol{q}}_i, \dot{\boldsymbol{q}}_i, \boldsymbol{q}_i], \tilde{\boldsymbol{\tau}}_i\}_{i=1}^m$ for the GPR.

As the continuity condition is fulfilled through assumption C3, the difference between the model error and the GP posterior tends to zero as the number of training points approaches infinity which incorporates that the variance is also zero. This allows the following conclusion:

Corollary 5. The tracking error for the rigid manipulator (4) with control law (5) is asymptotically stable in the large as the number of training points for the GP approaches infinity.

Proof. If the number of training points tends to infinity, lemma 1 shows that

$$\boldsymbol{\mu}(\boldsymbol{q_c}) = -\tilde{H}(\boldsymbol{q})\ddot{\boldsymbol{q}}_{\boldsymbol{d}} - \tilde{C}(\boldsymbol{q}, \dot{\boldsymbol{q}}_{\boldsymbol{d}})\dot{\boldsymbol{q}}_{\boldsymbol{d}} - \tilde{\boldsymbol{g}}(\boldsymbol{q})$$

holds. The upper bound for the drift operator of the Lyapunov function (11) can now be rewritten as

$$\begin{aligned} LV \leq & -\dot{\boldsymbol{e}}^\top (K_d - \varepsilon H)\dot{\boldsymbol{e}} - \varepsilon \boldsymbol{e}^\top K_p \boldsymbol{e} + \varepsilon \boldsymbol{e}^\top (\dot{H} - C - K_d)\dot{\boldsymbol{e}} \\ & + \|\dot{\boldsymbol{e}}\|\beta\|\dot{\boldsymbol{e}}\| + \varepsilon \boldsymbol{e}^\top \tilde{C}(\boldsymbol{q}, \dot{\boldsymbol{q}}_{\boldsymbol{d}})\dot{\boldsymbol{e}}. \end{aligned}$$

With C2, the tracking error is asymptotically stable in the large. Thus for the limit value consideration of infinite many training points, the tracking error will converge to zero for any $[\dot{\boldsymbol{q}}_0^\top, \boldsymbol{q}_0^\top]^\top \in \mathcal{X}$. ■

Remark 2
The GP is trained over $\ddot{\boldsymbol{q}}, \dot{\boldsymbol{q}}, \boldsymbol{q}$ but receives $\ddot{\boldsymbol{q}}_{\boldsymbol{d}}, \dot{\boldsymbol{q}}_{\boldsymbol{d}}, \boldsymbol{q}$ as inputs in the control law. This is beneficial for practical implementation as then no feedback of the manipulator's acceleration and the velocity is required. Additionally, the dependency of $\tilde{C}$ on $\dot{\boldsymbol{q}}$ is problematic because it cannot be identified isolated from the the angular velocity $\dot{\boldsymbol{q}}$ which is multiplied with $\tilde{C}$. Therefore, we use the desired velocity for the input of the Gaussian Process.

4. SIMULATION AND EXPERIMENT

4.1 Simulation

As unmanned aerial vehicles (UAVs) increasingly gain importance in automation and robotics, for simulation, we consider a model of a NACA-0015 airfoil moved through the air as illustrated in Fig. 3. The inertia J_a of the wing is assumed as $1\,\mathrm{kgm}^2$, the mass $m = 1\,\mathrm{kg}$ and the distance between the joint and the center of mass $l = 1\,\mathrm{m}$. The goal is to control the angle q of the wing with an input torque τ. The wing is affected by an aerodynamic force which can be decomposed in lift and drag.
These forces depend on the angle of attack which is the angle between the direction of the air flow and the reference line of the wing. For a large angle of attack the lift and drag force are highly nonlinear and difficult to model mathematically since air flow becomes turbulent. Our simulations are based on the measurements of wing taken in a wind tunnel [Sheldahl and Klimas, 1981]. For the model shown in Fig. 3, the lift and drag forces are converted in the resulting torque and gravity is added.

Assume a damping free pendulum for the estimated dynamics

$$\hat{J}_a \ddot{q} + \hat{m} g \hat{l} \sin(q) = \hat{\tau},$$

with the estimated parameters $\hat{J}_a = 0.9 J_a$, $\hat{l}\hat{m} = 0.9ml$. Figure 4 shows the simulation results for the classical augmented PD control law using the estimated model. The feedback terms are set to $K_p = K_d = 5$ and the desired

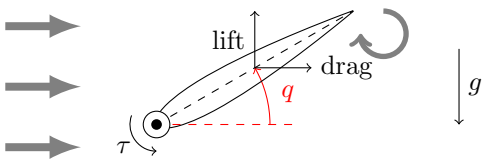

Fig. 3. Model of torque controlled wing. Lift / drag forces are highly nonlinear functions of the angle of attack q.

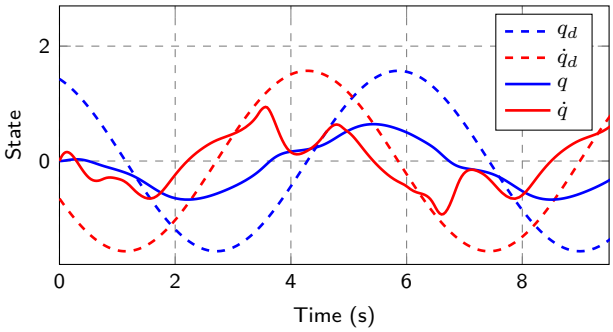

Fig. 4. Classical augmented PD control with an estimated model does not generate satisfactory results in comparison to the proposed control law which is shown in Fig. 5. The dashed lines are the desired joint position and velocity and the solid lines show the true values.

trajectory (dashed) is sinusoidal. Since the model contains parameter imprecision and influence of the airflow is not covered, the joint angle q differs from the desired q_d.

In the next step, the proposed control law (5) is used. First, the GP learns the difference between the estimated model and the real wing. For this purpose, we generate 990 homogeneous distributed pairs of torques τ and initial positions q_0 on the set $[-8, 8] \times [-\pi, \pi]$ to generate training points as shown in Fig. 2. The initial joint velocity and acceleration is set to zero. In this example, we do not use an extra controller but apply the torque directly for a short time interval to the manipulator. After 0.5 s the joint position, velocity and acceleration $\{\ddot{q}, \dot{q}, q\}$ are recorded. These values are inserted into the model to compute the estimated torque $\hat{\tau}$. The difference between the applied torque and the estimated torque $\tilde{\tau} = \tau - \hat{\tau}$ is saved. The values $\{\ddot{q}, \dot{q}, q\}$ and $\{\tilde{\tau}\}$ build up a training pair.

The GP is trained by this collection of training pairs and the hyperparameters of the squared exponential covariance function are optimized with a gradient method. Afterwards, the proposed control law (5) with the same desired trajectory and feedback gains is used. To show the effect of the stochastic control law, we simulate 1000 realizations of the stochastic differential equation with a sample time of 1 ms. Figure 5 shows the mean (solid) and standard deviation (gray area) of the joint angle/velocity and the desired angles/velocity (dashed). The stochastic behavior is based on the stochastic prediction of the GP since the finite number of training data generates only an uncertain model. Since the GP cancels the uncertainties of the model, the mean of the joint angles converges to a tight bound around the desired angles. The size of the standard deviation depends on the certainty of the prediction of the GP which is influenced by the number and the distribution of the training points.

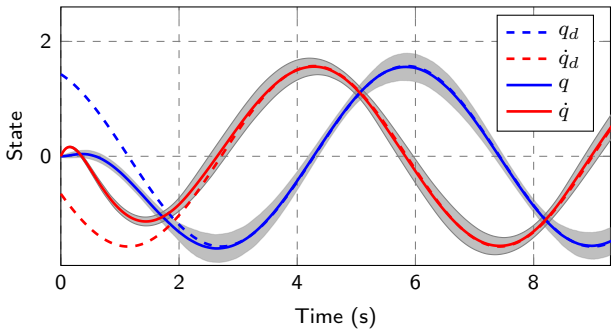

Fig. 5. The proposed GP-based control law strongly reduces the tracking error in comparison to the classical augmented PD control. The mean (solid line) of the joint angle/velocity converges to a tight bound around the desired trajectory (dashed line). The shaded area marks the 2σ interval of the 1000 simulations.

4.2 Experimental Evaluation

Setup For the experimental evaluation, we use the 3-dof SCARA robot CARBO as pictured in Fig. 6. The links between the joints have a length of 0.3 m. Since the third joint just rotates a camera which is mounted as end effector, this joint is fixed for the experiment. A low level PD-controller control enforces the generated torque by regulating the voltage based on a measurement of the current (which is approximately proportional to the torque). The robot manipulates a flexible rubber band which is fixed on the right side of the workspace. There exists no precise model for the flexible, nonlinear behavior of the rubber band, which makes the learning approach necessary. The task is for example comparable with the handling of rubber seals in the automotive manufacturing. The desired trajectory follows a sinusoidal shape with a frequency of $1\,\mathrm{s}^{-1}$ for the first, $2\,\mathrm{s}^{-1}$ for the second joint and an amplitude of $\pi/5$. The controller is implemented in MATLAB/Simulink on a Linux real-time system with a sample rate of 1 ms.

Task evaluation For the evaluation of the proposed method, we compare five different controllers on the same desired trajectory.

- HG-PD: A high gain PD controller with the parameters $K_P^{(HG)} = \mathrm{diag}(800, 600)$ and $K_D^{(HG)} = \mathrm{diag}(5, 5)$ without any feed forward model.
- LG-PD: A low gain PD controller with the parameters $K_P^{(LG)} = \mathrm{diag}(20, 15)$ and $K_D^{(LG)} = \mathrm{diag}(5, 5)$ without any feed forward model.
- CT: A computed-torque controller based on a friction free model of the robot which is generated from the CAD-model combined with the LG-PD.
- CT-SP: A computed-torque controller based on a friction free model of the robot and a linear model of the rubber band combined with the LG-PD.
- CT-GP: A modified computed-torque controller based on a friction free model of the robot and the trained GP (our approach) in combination with the LG-PD.

The high gain approach (HG-PD) is not directly comparable to the other approach as it suffers from many disadvantages as discussed in the introduction, but it serves as a "ground truth" here. It was also employed to generate

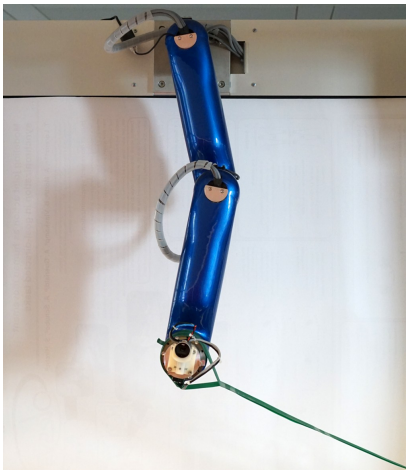

Fig. 6. A picture of the 3-dof robot CARBO with a rubber band between the robot's end effector and the ground.

the training data for the CT-GP approach by recording 351 training points corrupted by sensor noise at a rate of 30 ms while the robot follows the desired trajectory. The GPR is implemented with the GPML toolbox [Rasmussen and Williams, 2006]. The hyperparameter of the Gaussian Process are obtained through a gradient-based likelihood maximization. To obtain the best performance, we employ the deterministic version of our controller, thus using the GP's mean function. The performance is evaluated using the root mean square error (RMSE) between the desired and the real position of the joint angles for all controllers.

Results Figure 7 shows the RMSE in both joints for the different controllers. The low gain controller (LG-PD) performs very poorly, since no model knowledge is employed. This behavior is improved by adding computed torque (CT). Since the accuracy of the first joint increases a lot, the error of the second joint is slightly worse. If the influence of the rubber band is taken into account (CT-SP), the accuracy for both joints is improved. Our approach (CT-GP) with low gain feedback clearly outperforms all other approaches with low gain and is even competitive with the high gain controller.
The applied torque for the first joint is visualized in Fig. 8. The CT-GP generates a torque which is very similar to the high gain controller, while all others clearly differ. The in-

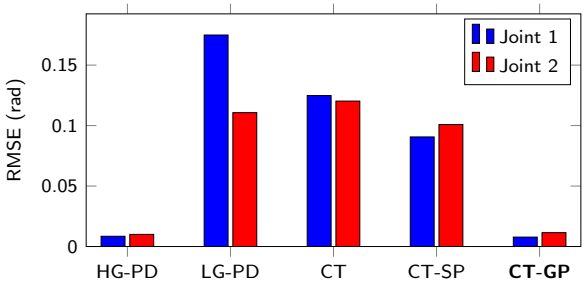

Fig. 7. The RMSE between desired and true joint angles for the different control laws. The error of the CT-GP is clearly smaller than for all other approaches with low gains. The high-gain approach (LG-PD) has similar RMSE but other undesired properties and therefore should not be directly compared.

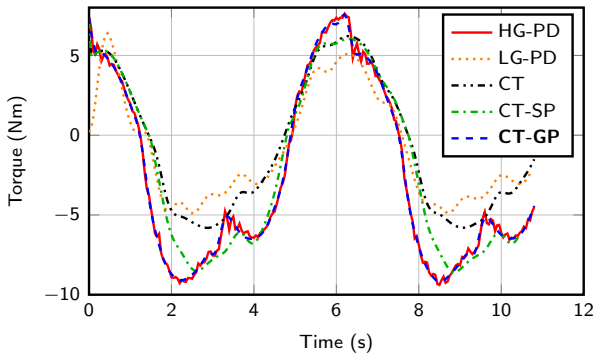

Fig. 8. Applied torque on first joint for different controllers.

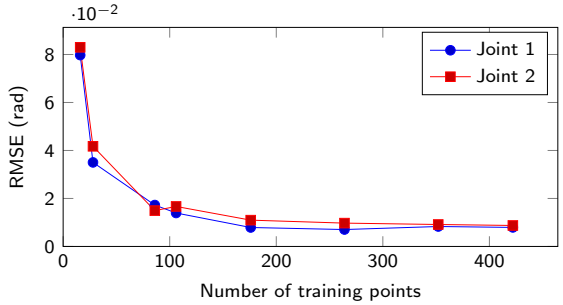

Fig. 9. The learning curve of the CT-GP with increasing number of training points.

fluence of the amount of training data on the performance of our approach is shown in Fig. 9. With an increasing number of training points, the error is decreasing.

Discussion The simulation and the experiment show that our approach does not only provide theoretical guarantees, but also shows performance advantages in real-world applications. The experiment showed that the feedback gains can be reduced by a factor of 40 while keeping the performance at a similar level. This was not achieved with the best analytically derived physical model for our scenario. The simulation showed how highly nonlinear effects (turbulent airflow) can also be captured by our nonparametric modeling approach and leads to guaranteed diminishing tracking error.

5. CONCLUSIONS

In this paper, we introduce a modified computed-torque control law based on Gaussian Process regression (GPR) for robotic manipulators. For this purpose, a GP learns the difference between an estimated model and the true robot. Afterwards, the control law uses the model and the GPR to compensate the robot dynamics. The derived method guarantees stochastic sample path boundedness around zero. If the number of training points tends to infinity, the tracking error becomes asymptotically stable. The proposed control law is of stochastic nature and the convergence occurs in probability in the large. Therefore, also its deterministic pendant (the GP's mean function) leads to the stable behavior.

REFERENCES

Alberto, N.T., Mistry, M., and Stulp, F. (2014). Computed torque control with variable gains through Gaussian process regression. In *International Conference on Humanoid Robots*, 212–217. IEEE-RAS.

Beckers, T. and Hirche, S. (2016). Equilibrium distributions and stability analysis of Gaussian process state space models. In *Conference on Decision and Control*, 6355–6361.

Deisenroth, M.P., Fox, D., and Rasmussen, C.E. (2015). Gaussian processes for data-efficient learning in robotics and control. *Transactions on Pattern Analysis and Machine Intelligence*, 37(2), 408–423.

Gard, T.C. (1988). *Introduction to Stochastic Differential Equations*. M. Dekker.

Ghorbel, F., Srinivasan, B., and Spong, M.W. (1993). On the positive definiteness and uniform boundedness of the inertia matrix of robot manipulators. In *Conference on Decision and Control*, 1103–1108. IEEE.

Mao, X. (2007). *Stochastic Differential Equations and Applications*. Elsevier.

Nguyen-Tuong, D. and Peters, J. (2008). Learning robot dynamics for computed torque control using local Gaussian processes regression. In *Symposium Learning and Adaptive Behaviors for Robotic Systems*, 59–64. IEEE.

Nguyen-Tuong, D. and Peters, J. (2010). Using model knowledge for learning inverse dynamics. In *International Conference on Robotics and Automation*, 2677–2682.

Rasmussen, C.E. and Williams, C.K. (2006). *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA, USA.

Sheldahl, R.E. and Klimas, P.C. (1981). Aerodynamic characteristics of seven symmetrical airfoil sections through 180-degree angle of attack for use in aerodynamic analysis of vertical axis wind turbines. Technical report, Sandia National Labs, Albuquerque, USA.

Siciliano, B., Sciavicco, L., Villani, L., and Oriolo, G. (2010). *Robotics: modelling, planning and control*. Springer Science+Business Media.

Slotine, J.J.E. and Li, W. (1987). On the adaptive control of robot manipulators. *The International Journal of Robotics Research*, 6(3), 49–59.

Spong, M.W. and Vidyasagar, M. (2008). *Robot dynamics and control*. John Wiley & Sons.

Sui, Y., Gotovos, A., Burdick, J., and Krause, A. (2015). Safe exploration for optimization with gaussian processes. In *Proceedings of The 32nd International Conference on Machine Learning*, 997–1005.

Takegaki, M. and Arimoto, S. (1981). A new feedback method for dynamic control of manipulators. *Journal of Dynamic Systems, Measurement, and Control*, 103(2), 119–125.

Vaart, A.v.d. and Zanten, H.v. (2011). Information rates of nonparametric Gaussian process methods. *Journal of Machine Learning Research*, 12, 2095–2119.

Whitcomb, L.L., Rizzi, A.A., and Koditschek, D.E. (1993). Comparative experiments with a new adaptive controller for robot arms. *IEEE Transactions on Robotics and Automation*, 9(1), 59–70.

Williams, C., Klanke, S., Vijayakumar, S., and Chai, K.M. (2009). Multi-task Gaussian process learning of robot inverse dynamics. In *Advances in Neural Information Processing Systems*, 265–272.