Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Gaussian Process Based Disturbance Compensation for an Inverted Pendulum

Harald Aschemann * **Cristina Tarín** **

* *Chair of Mechatronics, University of Rostock, Justus-von-Liebig Weg 6, 18059 Rostock, Germany (e-mail: Harald.Aschemann@uni-rostock.de).*
** *Institute for System Dynamics, University of Stuttgart, Waldburgstr. 17/19, 70563 Stuttgart, Germany (e-mail: cristina.tarin@isys.uni-stuttgart.de)*

Abstract: This paper considers an inverted pendulum with velocity input to demonstrate the validity of a new method for estimating unknown disturbances such as nonlinear friction and damping. With only one run of a given swing-up strategy, the input and output data of the system are collected. They are the basis for a Gaussian process modelling that is used to estimate the unknown disturbance term. The learned Gaussian process model can be used subsequently to predict this disturbance and serve for an online disturbance compensation. Simulation results and a comparison with a classical observer-based disturbance compensation indicate the benefits of the proposed approach.

Keywords: Gaussian Process, Extended Linearisation, State and Disturbance Observer, Disturbance Compensation, Inverted Pendulum.

1. INTRODUCTION

There exists a huge variety of mechanical systems such as flexible robots or manipulators, which are characterised by nonlinearities or under-actuation, which increase the difficulty to control these systems. Especially in the presence of unknown disturbances, the control task becomes challenging. In the past few years, increasing demands on new techniques for underactuated and nonlinear systems with unknown disturbances have been reported, see e.g. Huang (2019).

An inverted pendulum (IP) driven by a moving carriage has a simple structure with one control input, which is usually the force generated by an electric motor. The measured variables are typically the angle of the pendulum and the position of the carriage, where the pendulum lacks of a drive and represents the under-actuated system part. Regarding the equations of motions of the pendulum, several terms depending on cosine and sine functions of the pendulum angle render it nonlinear. Especially during the swing-up process – with a wide range of the pendulum angle and high angular velocities – the control and observer design becomes challenging. Thus, the inverted pendulum has become a benchmark system for nonlinear underactuated systems as also shown in Tian (2018). The results presented in the given contribution can be transferred to more complex robotic systems with similar characteristics.

A large variety of modelling approaches and balancing control laws have already been developed for the IP system, see e.g. Huang (2020) and Yang (2018). For modelling, first principle models are frequently used. However, an accurate and reliable first principle model is often not achievable due to nonlinear friction characteristics or damping effects, which are difficult to model. Static and dynamic friction models are known in the literature but still face certain shortcomings and require additional experiments for an appropriate parametrisation.

As for robot and manipulator control, various approaches have been proposed in the last decades like computed-torque control, see Pekarovskiy (2018), as a special case of feedback linearisation, where the nonlinear system is transformed into an equivalent linear system through a change of variables and a suitable control input. When an exact system model is available, this type of control law is able to compensate for the nonlinear robot dynamics to enable an accurate tracking control. Nevertheless, the success of this approach always strongly depends on the model precision.

Other approaches to feedback control of nonlinear underactuated systems include the extended linearisation method for systems that can be described in a quasi-linear form with state and parameter-dependent matrices. For this class of systems, the gain matrix of the state feedback controllers may be determined by an eigenvalue placement as reported in Rugh (1986) and Rauh (2017) or by state-dependent Riccati equations (SDREs). However, there is still an important problem to deal with which is caused by unknown disturbances affecting the system behaviour. In fact, the disturbance present in case of the IP benchmark system is not measurable. Therefore, it has to be estimated in order to allow for the design of a disturbance compensation. Observer-based estimation involves the design of a combined state and disturbance observer.

On the other hand, when tackling problems by data-driven approaches, Bayesian modelling leads to generative models optimized to fit measurement data under explicit assumptions w.r.t. measurement uncertainty. Gaussian processes (GPs) are highlighted as an important class of Bayesian nonparametric models. However, GP-related, data-driven control methods have remained so far, with notable exceptions as presented in Ringkowsky (2019), Zonghai (2014) and Beckers (2017), largely unexploited by the general control engineering community.

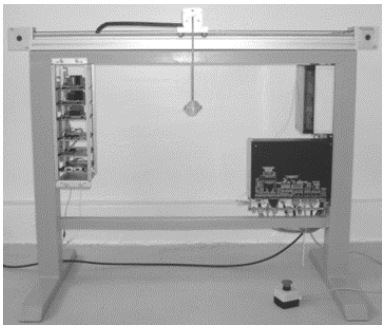

Fig. 1. Photo of the inverted pendulum test rig.

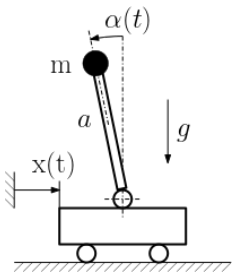

Fig. 2. Mechanical model of the inverted pendulum.

Whereas directly measurable quantities are often modelled by GPs as presented in Chowdhary (2013), Ko (2009) and Deisenroth (2015), in this paper the focus is on modelling of unknown friction and damping characteristics that are not directly accessible by measurements. The main idea is to employ a state and disturbance observer in a learning phase to reconstruct the variable of interest, determine a GP representation and to use this GP model subsequently for an online compensation.

The remainder of the paper is structured as follows: In Sect. 2, the mathematical model of an inverted pendulum with a velocity input is derived. As pointed out before, this model is not the standard one. Sect. 3 is devoted to the control design using extended linearisation. The observer-based compensation of external disturbances of Sect. 4 is compared to the disturbance estimation based on a Gaussian process modelling as discussed in Sect. 5. Corresponding simulation results are shown and a comparative analysis between these two approaches is performed in Sect. 6. Finally, in Sect. 7, the main findings and outcomes of this paper are discussed briefly.

2. SYSTEM DYNAMICS OF AN INVERTED PENDULUM WITH VELOCITY INPUT

In the following, a nonlinear model for an inverted pendulum with an underlying velocity control of the carriage is presented. The corresponding test rig at the Chair of Mechatronics, University of Rostock, is depicted in Fig. 1.

The carriage is driven by a DC servo motor via a toothed belt. The horizontal position $x(t)$ of the carriage can be derived from the rotor angle measurement, accessible at the control unit at the current converter. The pendulum consists of a rod (length a) with negligible mass and an end mass (mass m) at the tip of the rod. The pendulum angle w.r.t. the upright position, which is measured by an optical encoder, is denoted by $\alpha(t)$, see Fig. 2. The equation of motion for the pendulum follows directly from a balance of angular momentum – in a form that considers an accelerated rotary joint

$$ma^2\ddot{\alpha}(t) = ma\cos(\alpha(t))\ddot{x}(t) + mga\sin(\alpha(t)) - \tau(t)\ , \quad (1)$$

where g stands for the gravitational acceleration, whereas $\tau(t)$ is an unknown term accounting for damping and nonlinear friction. In the simulation, this torque is represented by

$$\tau(t) = b\dot{\alpha}(t) + m_{R0}\tanh\left[\frac{\dot{\alpha}(t)}{\varepsilon}\right]\ ,\ \varepsilon << 1. \quad (2)$$

Here, m_{R0} characterises the amplitude of the nonlinear friction term, whereas ε defines its steepness at zero angular velocity. The motion of the carriage is characterized by a first-order lag behaviour of an underlying velocity control loop

$$T_1\ddot{x}(t) + \dot{x}(t) = u(t)\ , \quad (3)$$

where T_1 represents the time constant. The controlled output is the horizontal position of the end mass

$$y(t) = x(t) - a\sin(\alpha(t))\ . \quad (4)$$

By choosing the state vector $\mathbf{x}(t) = [\alpha(t)\ \dot{\alpha}(t)\ x(t)\ \dot{x}(t)]^T$, the nonlinear state-space representation becomes

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \dot{\alpha}(t) \\ \frac{g}{a}\sin(\alpha(t)) + \frac{1}{aT_1}[u(t) - \dot{x}(t)]\cos(\alpha(t)) + \tau_R(t) \\ \dot{x}(t) \\ \frac{1}{T_1}[u(t) - \dot{x}(t)] \end{bmatrix}\ , \quad (5)$$

with the unknown acceleration $\tau_R(t) = -\frac{\tau(t)}{ma^2}$ that is proportional to the torque $\tau(t)$.

3. CONTROL DESIGN USING EXTENDED LINEARISATION

For the control design, the nonlinear dynamic system model is reformulated and rewritten in a quasi-linear state-space representation

$$\begin{bmatrix} \dot{\alpha}(t) \\ \ddot{\alpha}(t) \\ \dot{x}(t) \\ \ddot{x}(t) \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{g}{a}\mathrm{si}(\alpha(t)) & 0 & 0 & -\frac{\cos(\alpha(t))}{aT_1} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -\frac{1}{T_1} \end{bmatrix}}_{\mathbf{A}(\alpha)} \cdot \underbrace{\begin{bmatrix} \alpha(t) \\ \dot{\alpha}(t) \\ x(t) \\ \dot{x}(t) \end{bmatrix}}_{\mathbf{x}(t)} \quad (6)$$

$$+ \underbrace{\begin{bmatrix} 0 \\ \frac{\cos(\alpha(t))}{aT_1} \\ 0 \\ \frac{1}{T_1} \end{bmatrix}}_{\mathbf{b}(\alpha)} \cdot u(t) + \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{e}} \cdot \tau_R(t) \quad (7)$$

with the function $\mathrm{si}(\alpha) = \sin(\alpha)/\alpha$, where $\mathrm{si}(0) = 1$ holds. In a similar manner, the output equation is written in quasi-linear form. This results in

$$y(t) = \underbrace{[-a\ \mathrm{si}(\alpha(t))\ 0\ 1\ 0]}_{\mathbf{c}^T(\alpha)} \cdot \mathbf{x}(t) \quad (8)$$

It becomes obvious that the system matrix and the input as well as output vector depend on the state variable $\alpha(t)$

$$\dot{\mathbf{x}}(t) = \mathbf{A}(\alpha)\mathbf{x}(t) + \mathbf{b}(\alpha)u(t) + \mathbf{e}\ \tau_R(t) \quad (9)$$

$$y(t) = \mathbf{c}^T(\alpha)\mathbf{x}(t) \quad (10)$$

Note that this state-dependent formulation is still exact and does not involve any approximations. In the range of $|\alpha| < \pi/2$, the system point-wise fulfils Kalman's controllability criterion, which guarantees a successful control design in the framework of extended linearisation. Here, a full rank of the state-dependent controllability matrix

$$\mathbf{Q}_S = \left[\mathbf{b}\ \mathbf{Ab}\ \mathbf{A}^2\mathbf{b}\ \mathbf{A}^3\mathbf{b}\right] \quad (11)$$

is obtained according to

$$\det \mathbf{Q}_S(\alpha) \neq 0 : \cos(\alpha)\mathrm{si}(\alpha) \neq 0\,. \tag{12}$$

3.1 Eigenvalue Assignment Using Extended Linearisation

Based on the state-dependent state-space representation, an eigenvalue assignment is performed. In the case of single-input single-output systems, a symbolic computation of a state-dependent gain vector $\mathbf{k}^T(\mathbf{x})$ is possible by a comparison of the characteristic equation

$$p(s) = \det\left(s\mathbf{I} - \mathbf{A}(\alpha) + \mathbf{b}(\alpha)\mathbf{k}^T(\alpha)\right) \tag{13}$$

with a desired characteristic polynomial of the closed-loop system, for example, by specifying a fourfold eigenvalue s_R

$$p_R(s) = (s - s_R)^4\,. \tag{14}$$

The control gain vector

$$\mathbf{k}^T = [k_{R1}\ \ k_{R2}\ \ k_{R3}\ \ k_{R4}] \tag{15}$$

is not constant but depends on the angle α, with

$$\begin{aligned}
k_{R1} &= \frac{T_1\left(a^2 {s_R}^4 + 6\,\mathrm{si}(\alpha)\,g\,a\,{s_R}^2 + (\mathrm{si}(\alpha))^2 g^2\right)}{g\,\mathrm{si}(\alpha)\cos(\alpha)},\\
k_{R2} &= -4\,\frac{aT_1 s_R\left(a\,s_R^2 + g\,\mathrm{si}(\alpha)\right)}{g\,\mathrm{si}(\alpha)\cos(\alpha)},\\
k_{R3} &= -\frac{s_R^4 T_1 a}{g\,si(\alpha)},\ k_{R4} = -\frac{-4\,s_R^3 T_1 a + g\,\mathrm{si}(\alpha)}{g\,\mathrm{si}(\alpha)}.
\end{aligned} \tag{16}$$

In the admissible range of $|\alpha| < \pi/2$, which can be kept during control operation by a corresponding trajectory planning, any singularities occur.

The stability of the time-varying system can be shown employing a quadratic Lyapunov function

$$V(\mathbf{x}) = \mathbf{x}^T \mathbf{P}\,\mathbf{x} \tag{17}$$

with a constant matrix $\mathbf{P} = \mathbf{I}$. A sufficient condition for asymptotic stability is the negative definiteness of the matrix

$$\mathbf{A}_R + \mathbf{A}_R^T < 0\,, \tag{18}$$

where $\mathbf{A}_R = \mathbf{A} - \mathbf{b}\,\mathbf{k}^T$ denotes the closed-loop system matrix. Alternatively, methods based on linear matrix inequalities (LMIs) could be employed to assess the closed-loop stability.

3.2 Feedforward Control Design Using Extended Linearisation

For feedforward control design, the horizontal position of the center of gravity (CoG) of the pendulum is chosen as the controlled output. The corresponding output equation results in

$$y(t) = [\,-a\,\mathrm{si}(\alpha(t))\ \ 0\ \ 1\ \ 0\,]\,\mathbf{x}(t) = \mathbf{c}^T(\alpha(t))\,\mathbf{x}(t)\,. \tag{19}$$

Then, the command transfer function becomes

$$\frac{Y(s)}{U_{FF}(s)} = \mathbf{c}^T\left(s\mathbf{I} - \mathbf{A} + \mathbf{b}\mathbf{k}^T\right)^{-1}\mathbf{b} = \frac{\left(b_0 + b_1 \cdot sb + b_2 \cdot s^2\right)}{N(s)}\,. \tag{20}$$

It becomes obvious that the numerator is given by a second-order polynomial in s with two transfer zeros. The chosen output, hence, is a non-flat output variable, which renders the feedforward control design more involved and demands for an alternative solution towards an appropriate feedforward control action u_{FF}. The main idea consists in modifying the numerator of the command transfer function by means of a polynomial ansatz in the Laplace domain

$$U_{FF}(s) = \left[k_{V0} + k_{V1} \cdot s + k_{V2} \cdot s^2 + k_{V3} \cdot s^3\right] Y_d(s)\,. \tag{21}$$

As the desired trajectory $y_{Kd}(t)$ as well as its first three time derivatives are provided by a trajectory planning module, this feedforward action is realizable. The parameter-dependent feedforward gains $k_{Vj} = k_{Vj}(\alpha)$ follow from a comparison of the corresponding coefficients in the numerator as well as the denominator polynomials of

$$\begin{aligned}
\frac{Y(s)}{Y_d(s)} &= \frac{\left(b_0 + b_1 \cdot s + b_2 \cdot s^2\right)\left[k_{V0} + \ldots + k_{V3} \cdot s^3\right]}{N(s)}\\
&= \frac{b_{V0}(k_{Vj}) + b_{V1}(k_{Vj}) \cdot s + \ldots + b_{V5}(k_{Vj}) \cdot s^5}{a_0 + a_1 \cdot s + a_2 \cdot s^2 + a_3 \cdot s^3 + s^4}
\end{aligned} \tag{22}$$

according to

$$a_i = b_{Vi}(k_{Vj}), i = 0, \ldots, n = 3\,. \tag{23}$$

In the given case, the feedforward gains become

$$\begin{aligned}
&k_{V0} = k_{R3},\ k_{V1} = k_{R4} + 1,\\
&k_{V2} = -\frac{\mathrm{si}(\alpha)\cos(\alpha)\,k_{R3}\,a - g\,\mathrm{si}(\alpha)\,T_1 + \cos(\alpha)\,k_{R1}}{g\,\mathrm{si}(\alpha)},\\
&k_{V3} = -\frac{\cos(\alpha)\left(\mathrm{si}(\alpha)\,k_{R4}\,a + a\,\mathrm{si}(\alpha) + k_{R2}\right)}{g\,\mathrm{si}(\alpha)}.
\end{aligned} \tag{24}$$

Obviously— caused by the higher numerator degree in (22) – perfect tracking is not achievable and remaining dynamics must be accepted. Nevertheless, this feedforward control is easily implementable and significantly improves the tracking behaviour. The gain-scheduling is performed using the measured angle $\alpha(t)$ in this work; alternatively, also desired values corresponding to the chosen trajectory could be employed instead.

Moreover, it allows for an analysis of the feasibility of desired state and output trajectories. If such constraints are violated, a proper time scaling of the desired trajectory could be applied to render it feasible.

4. OBSERVER-BASED COMPENSATION OF EXTERNAL DISTURBANCES

Disturbance behaviour and tracking accuracy in view of nonlinear friction and model uncertainty can be improved significantly by the introduction of a compensating control action using an appropriate estimator.

4.1 Design of a Combined State and Disturbance Observer Using Extended Linearisation

As only the carriage position $x(t)$ and the pendulum angle $\alpha(t)$ are measurable, a combined state and disturbance observer is designed. By extending the state vector by the disturbance $\tau_R(t)$, the measurement equation becomes

$$\mathbf{y}(t) = \begin{bmatrix} x(t) \\ \alpha(t) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}}_{\mathbf{C}} \cdot \underbrace{\begin{bmatrix} \alpha(t) \\ \dot{\alpha}(t) \\ x(t) \\ \dot{x}(t) \\ \tau_R(t) \end{bmatrix}}_{\mathbf{x}_O}. \tag{25}$$

As disturbance model for the disturbance $\tau_R(t)$, a single integrator is appropriate: $\dot{\tau}_R(t) = 0$. The ansatz for a combined state and disturbance observer is given by the following state-space representation

$$\begin{bmatrix}\dot{\hat{\alpha}}(t)\\ \ddot{\hat{\alpha}}(t)\\ \dot{\hat{x}}(t)\\ \ddot{\hat{x}}(t)\\ \dot{\hat{\tau}}_R(t)\end{bmatrix} = \underbrace{\begin{bmatrix}0 & 1 & 0 & 0 & 0\\ \frac{g}{a}\mathrm{si}(\alpha(t)) & 0 & 0 & -\frac{\cos(\alpha(t))}{aT_1} & 1\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -\frac{1}{T_1} & 0\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}}_{\mathbf{A}_O} \cdot \underbrace{\begin{bmatrix}\hat{\alpha}(t)\\ \dot{\hat{\alpha}}(t)\\ \hat{x}(t)\\ \dot{\hat{x}}(t)\\ \hat{\tau}_R(t)\end{bmatrix}}_{\hat{\mathbf{x}}_O} + \underbrace{\begin{bmatrix}0\\ \frac{\cos(\alpha(t))}{aT_1}\\ 0\\ \frac{1}{T_1}\\ 0\end{bmatrix}}_{\mathbf{b}_O} \cdot u(t) + \underbrace{\begin{bmatrix}h_{11} & 0\\ h_{21} & 0\\ 0 & h_{32}\\ 0 & h_{42}\\ h_{51} & 0\end{bmatrix}}_{\mathbf{H}_O} \cdot \begin{bmatrix}x(t)-\hat{x}(t)\\ \alpha(t)-\hat{\alpha}(t)\end{bmatrix} \quad (26)$$

and yields an estimate of the disturbance $\hat{\tau}_R(t)$. The input variable is given by the desired carriage velocity $u(t)$ provided by the overall control law. Resulting from the choice for the observer gain matrix $\mathbf{H}_O$ above, the corresponding gains can be calculated in symbolic form. This becomes possible by a comparison of the characteristic equation

$$p(s) = \det\left(s\mathbf{I} - \mathbf{A}_O(\alpha) + \mathbf{H}_O(\alpha)\,\mathbf{C}\right) \quad (27)$$

with a desired characteristic polynomial of the observer, e.g., by specifying a five-fold real eigenvalue $s_O < 0$

$$p_O(s) = (s - s_0)^5 \ . \quad (28)$$

The observer gains depend on the angle $\alpha(t)$ and result in

$$h_{11} = 3\,s_O,\ h_{21} = \frac{3\,a\,s_O^2 + g\,\mathrm{si}(\alpha)}{a}$$
$$h_{32} = \frac{2\,s_O\,T_1 - 1}{T_1},\ h_{42} = \frac{T_1^2 s_O^2 - 2\,s_O\,T_1 + 1}{T_1^2},\ h_{51} = s_O^3\ . \quad (29)$$

The stability of the the closed-loop observer system matrix

$$\mathbf{A}_{O,cl} = \begin{bmatrix}-3\,s_O & 1 & 0 & 0 & 0\\ -3\,s_O^2 & 0 & 0 & -\frac{\cos(\alpha)}{a\,T_1} & 1\\ 0 & 0 & -\frac{2\,s_O\,T_1-1}{T_1} & 1 & 0\\ 0 & 0 & -\frac{T_1^2 s_O^2 - 2\,s_O\,T_1 + 1}{T_1^2} & -\frac{1}{T_1} & 0\\ -s_O^3 & 0 & 0 & 0 & 0\end{bmatrix}, \quad (30)$$

where only the $(2,4)$-element depends on α, has been investigated by Lyapunov methods by determining a joint Lyapunov function

$$V(\mathbf{x}) = \mathbf{x}^T \mathbf{P}_O\,\mathbf{x}\,, \quad (31)$$

with a constant matrix $\mathbf{P}_0 > 0$, for the given range of α. Note that this observer is applicable also during the envisaged swing-up control.

4.2 Model-Based Disturbance Compensation

The compensation is derived from the command transfer function

$$G_b(s) = \mathbf{c}^T\left(s\mathbf{I} - \mathbf{A} + \mathbf{b}\mathbf{k}^T\right)^{-1}\mathbf{b} \quad (32)$$

as well as the disturbance transfer function

$$G_e(s) = \mathbf{c}^T\left(s\mathbf{I} - \mathbf{A} + \mathbf{b}\mathbf{k}^T\right)^{-1}\mathbf{e} \quad (33)$$

to the controlled output. For disturbance compensation, the following design condition must be fulfilled in the Laplace domain up to linear terms in s

$$G_b(s)\cdot(k_{S0} + k_{S1}\cdot s) + G_e(s) = 0. \quad (34)$$

This leads to the following compensation gains

$$k_{S0} = -\frac{a\,(\mathrm{si}\,(\alpha)\,a\,k_{R3} + k_{R1})}{g\,\mathrm{si}\,(\alpha)}, \quad (35)$$

$$k_{S1} = -\frac{a\,(\mathrm{si}\,(\alpha)\,a\,k_{R4} + a\,\mathrm{si}\,(\alpha) + k_{R2})}{g\,\mathrm{si}\,(\alpha)}\ . \quad (36)$$

Disturbance compensation is performed by introducing the estimated force $\hat{\tau}_R(t)$ into the dynamic compensation law

$$u_{DC}(t) = k_{S0}\,\hat{\tau}_R(t) + k_{S1}\,\dot{\hat{\tau}}_R(t). \quad (37)$$

The necessary time derivative of the disturbance can be obtained approximately by a low-pass-filtered differentiation of $\hat{\tau}_R(t)$.

5. DISTURBANCE ESTIMATION THROUGH GAUSSIAN PROCESS MODELLING

An alternative to the observer-based estimation of the disturbance term as presented in Sect. 4 is given by a data-based approach. In this section, hence, some general ideas regarding data-driven modelling based on Gaussian Processes (GP) and the estimation of the disturbance using this GP modelling are presented.

5.1 Gaussian Process Regression

In a formal way, a GP can be defined as a collection of random variables (RV) with joint Gaussian distributions for any finite subset of RVs. The GP is, hence, fully described by a mean function and covariance or kernel function $K(\cdot,\cdot)$. The task consists in an approximation of a nonlinear mapping of a system

$$y = f(\mathbf{z}) + \varepsilon \quad (38)$$

from the input vector $\mathbf{z}$ in the input space $\mathbb{R}^{n_z}$ to a scalar output y in the output space $\mathbb{R}$, where white additive Gaussian noise $\varepsilon \sim \mathcal{N}(0,\sigma^2)$ characterises the observation noise. Considering an a-posteriori probability distribution that is derived from the training data set, the mean value corresponds to the estimate itself, whereas the variance represents an uncertainty measure of the estimate. The training data set consists of l training points and is defined as $\mathbf{Z} = [\mathbf{z}_1,\ldots,\mathbf{z}_l]^T \in \mathbb{R}^{l\times n_z}$. Then, the output data is a normal distribution

$$y \sim \mathcal{N}(0, \mathscr{K} + \sigma_\varepsilon^2 I)\,, \quad (39)$$

where the $l\times l$ covariance matrix can be constructed using the kernel function $[\mathscr{K}]_{ij} = K(\mathbf{Z}_i,\mathbf{Z}_j), i,j = 1,\ldots,l$ and the second term increases the noise variance on the main diagonal.

GP may be classified as an essential kernel machine learning method describing the relationship between input and output data by a kernel function $K(\cdot,\cdot)$. In general, the kernel function, i.e. covariance function, can be any function – provided that it generates a positive definite covariance matrix. The common choice, however, of the kernel function is the Gaussian function, which is defined as

$$K(\mathbf{Z}_i,\mathbf{Z}_j) = \sigma_f^2 \cdot \exp\left(-\frac{1}{2}(\mathbf{Z}_i - \mathbf{Z}_j)^T\Lambda^{-1}(\mathbf{Z}_i - \mathbf{Z}_j)\right), \quad (40)$$

where $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_l)$ denotes the characteristic length scale and σ_f is the signal standard deviation. These hyper-parameters – the length scales λ_i and the variance σ_f^2 – should be learned. For the GP, once the hyper-parameters have been learned, the mean and the variance of an output estimate (or prediction) is computed through the Gaussian joint probability

distribution function of the training data $\mathbf{Z}$ and the prediction data $\mathbf{z}^*$ as follows

$$\begin{aligned} \mu_{GP} &= K(\mathbf{Z},\mathbf{z}^*)^T(\mathscr{K}+\sigma_\varepsilon^2 I)^{-1}\mathbf{y}, \\ \sigma_{GP}^2 &= K(\mathbf{z}^*,\mathbf{z}^*)-K(\mathbf{Z},\mathbf{z}^*)^T(\mathscr{K}+\sigma_\varepsilon^2 I)^{-1}K(\mathbf{Z},\mathbf{z}^*). \end{aligned} \tag{41}$$

Therefore, the mean μ_{GP} can be used as an estimate for the corresponding output and the variance σ_{GP}^2 as an uncertainty measure of the output.

5.2 Energy-Based Swing-Up Control

The learning of the GP model takes place during an energy-based swing-up control of the pendulum from the stable equilibrium at $\alpha(0) = -\pi, \dot{\alpha}(0) = 0$. The design follows the ideas presented in Zhong (2001). For this purpose, the total mechanical energy of the pendulum subsystem is considered

$$E(\alpha(t),\dot{\alpha}(t)) = \frac{m\,a^2}{2}\dot{\alpha}^2(t) + m\,g\,a\,cos(\alpha(t)) \,. \tag{42}$$

The time derivative results in

$$\dot{E}(\alpha(t),\dot{\alpha}(t)) = m\,g\,a\,cos(\alpha(t))\,\dot{\alpha}(t)\,\upsilon(t) \,, \tag{43}$$

where $\upsilon(t) = \ddot{x}(t)$ denotes the carriage acceleration. With the energy $E_{up} = m\,g\,a$ in the upper pendulum position and the error $\tilde{E}(\alpha,\dot{\alpha}) = E(\alpha,\dot{\alpha}) - E_{up}$, an energy-based control law can be stated that brings the pendulum into the upright position

$$\upsilon(t) = -\tilde{k}\cdot\tilde{E}(t)\cdot[m\,g\,a\,cos(\alpha(t))\,\dot{\alpha}(t)]^2 \,, \tag{44}$$

with a positive constant $\tilde{k} > 0$. The physical control input follows from

$$u(t) = T_1\,\upsilon(t) + \dot{x}(t) \,. \tag{45}$$

If a predefined band around the upright position is attained, the control is switched to the gain-scheduled tracking control law described in Sect. 3.

6. SIMULATION RESULTS

In the presented simulation study, the main freedom in the design of the GP model is the definition of both type and number of inputs characterizing the model structure. For the modelling of the unmeasurable disturbance $\tau_R(t)$ with GP regression, an observer-based approach is employed in this work. The state and disturbance observer provides estimates for the angular velocity $\hat{\dot{\alpha}}(t)$ of the pendulum as well as the corresponding disturbance $\hat{\tau}_R(t)$. These estimates serve as inputs of a GP regression as described above.

For the training phase, the swing-up control according to Subsect. 5.2 is activated, and the GP regression is performed. To obtain realistic results, the measurements involve additional white Gaussian noise with a standard deviation of $\sigma_{noise} = 1e^{-6}$. The carriage position and the pendulum angle during the swing-up phase are illustrated by Fig. 3.

After the GP has been trained, the previously used observer-based disturbance compensation is deactivated and now substituted by an online prediction of the trained GP disturbance model.

Fig. 4 shows the simulated damping and friction term $\tau_R(\dot{\alpha})$, which is considered as unknown, and the learned GP regression model. Obviously, the characteristic w.r.t. the angular velocity of the pendulum is well reconstructed and may be used for a subsequent disturbance compensation. Fig. 5 depicts the time behaviour of the simulated damping and friction term, the estimated disturbance of the observer as well as the predictions

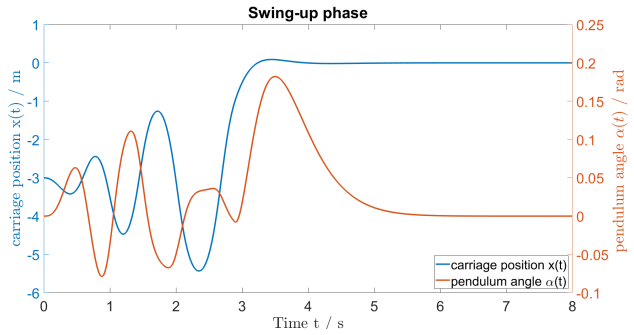

Fig. 3. Energy-based swing-up motion of the inverted pendulum: carriage position $x(t)$ and pendulum angle $\alpha(t)$.

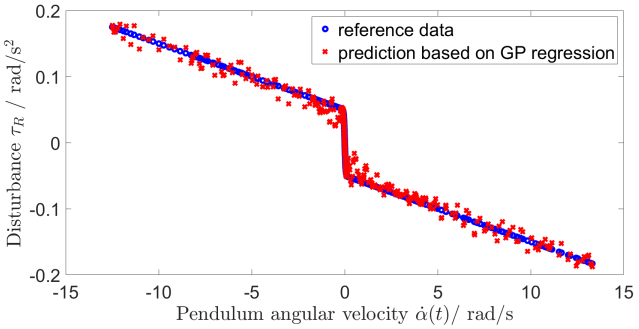

Fig. 4. Simulated damping and friction term $\tau_R(t)$ in comparison with the learned GP regression model.

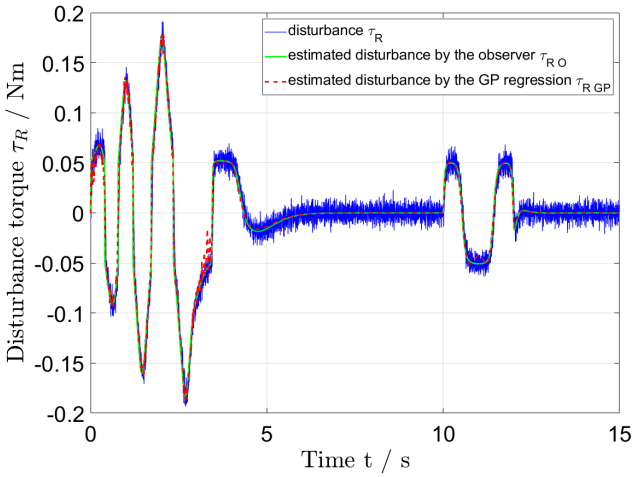

Fig. 5. Comparison of the simulated damping and friction term, the estimated disturbance of the observer and predictions from the learned GP regression model over time.

form the learned GP regression model. It can be seen that the estimate of the observer $\tau_{R,O}(t)$ is quite noisy, whereas both the simulated term $\tau_R(t)$ and the GP predictions $\tau_{R,GP}(t)$ are significantly smoother. After the successful learning of a GP regression model, the desired trajectory shown in Fig. 6 should be tracked accurately. For this purpose, a disturbance compensation is beneficial. The simulated values depicted here correspond to the GP-based disturbance compensation. Fig. 7 shows the benefits of a disturbance compensation using the learned GP model instead of a classical observer-based disturbance compensation. The tracking errors w.r.t. the desired trajectory

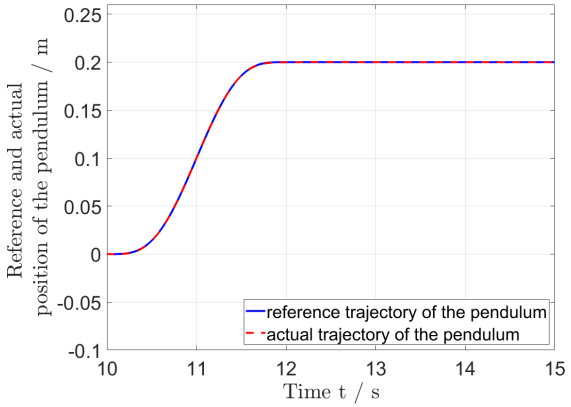

Fig. 6. Comparison of desired and simulated values for horizontal position $y(t)$ of the pendulum center of gravity according to (19) as the controlled output.

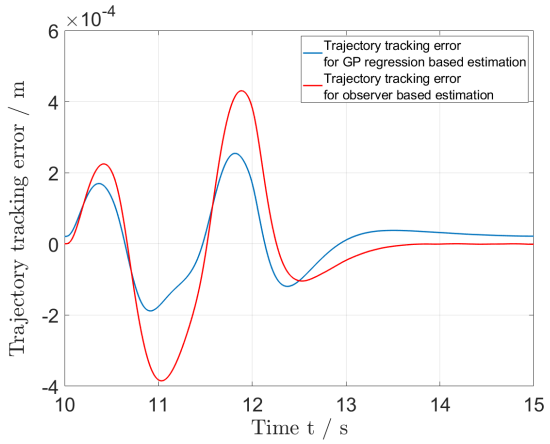

Fig. 7. Tracking errors w.r.t. a desired trajectory in sidewards direction: comparison of an observer-based disturbance compensation and a compensation using the GP regression model.

in sidewards direction indicate a reduction of the maximum tracking error by approx. 40 % that could be achieved.

7. CONCLUSIONS

In this paper, a nonlinear observer-based tracking control for an inverted pendulum with an underlying velocity control is presented. The fourth-order system model consists of the pendulum dynamics and the dynamics of the carriage motion, which is covered by a first-order lag behaviour for the carriage velocity and the kinematic relationship between position and velocity. The system model is characterized, moreover, by a nonlinear disturbance torque in the rotary joint of the pendulum which is compensated by GP regression techniques. The GP regression is performed using the estimated angular velocity and the estimate for the disturbance term provided from the observer during a swing-up control. Afterwards, the learned GP model is used for a disturbance compensation to improve the tracking of a desired trajectory in sidewards direction. The nonlinear control design is performed with extended linearisation techniques and involve a combination of dynamic feedforward control and feedback control. Stability of the closed-loop control loop as well as of the observer dynamics has been investigated by Lyapunov methods. Simulation results show clearly that the disturbance compensation using the GP regression model outperforms the observer-based disturbance compensation.

REFERENCES

J. Huang, S. Ri, T. Fukuda and Y. Wang *A Disturbance Observer Based Sliding Mode Control for a Class of Underactuated Robotic System With Mismatched Uncertainties.* In IEEE Transactions on Automatic Control, vol. 64, no. 6, 2019.

X. Tian and H. Peng. *A Model Predictive Control Approach to Inverted Pendulum System Based on RBF-ARX Model*. In Proceedings of the 37th Chinese Control Conference July 25-27, Wuhan, China, 2018.

J. Huang, M. Zhang, S. Ri, C. Xiong, Z. Li and Y. Kang. *High-Order Disturbance-Observer-Based Sliding Mode Control for Mobile Wheeled Inverted Pendulum Systems*. In IEEE Transactions on Industrial Electronics, vol. 67, no. 3, 2020.

X. Yang and X. Zheng *Swing-Up and Stabilization Control Design for an Underactuated Rotary Inverted Pendulum System: Theory and Experiments*. In IEEE Transactions on Industrial Electronics, vol. 65, no. 9, 2018.

W. J. Rugh. *An Extended Linearization Approach to Nonlinear System Inversion* in Transactions on Automatic Control, vol. AC-31. no. 8, August 1986.

A. Rauh, J. Prabel and H. Aschemann. *Oscillation Attenuation for Crane Payloads by Controlling the Rope Length Using Extended Linearization Techniques*. Proceedings of the 22nd International Conference on Methods and Models in Automation and Robotics (MMAR), Miedzyzdroje, Poland, 2017.

M. Ringkowski and O. Sawodny. *Gaussian Process Based Multi-Rate Observer for the Dynamic Positioning Error of a Measuring Machine*. In 2019 18th European Control Conference (ECC) Napoli, Italy, June 25-28, 2019.

S. Zonghai. *Gaussian Process Adaptive Control of Nonlinear System Based on Online Algorithm*. Proceedings of the 33rd Chinese Control Conference July 28-30, Nanjing, China, 2014.

T. Beckers, J. Umlauft and S. Hirche. *Stable Model-based Control with Gaussian Process Regression for Robot Manipulators.* In Proceedings of the 20th IFAC World Congress, Toulouse, France, 2017.

G. Chowdhary, H. A. Kingravi, J. P. How, and P. A. Vela. *A Bayesian Nonparametric Approach to Adaptive Control Using Gaussian Processes.* In Proc. IEEE 52nd Annu. Conf. Decision and Control, pp. 874-879, 2013.

W. Zhong and H. Röck. *Energy and Passivity Based Control of the Double Inverted Pendulum on a Cart.* In Proc. IEEE Conf. Control Applications, pp. 896-901, 2001.

J. Ko and D. Fox. *GP-Bayes filters: Bayesian Filtering Using Gaussian Process Prediction and Observation Models.* In Auton. Robots, vol. 27, no. 1, pp. 75-90, 2009.

M. P. Deisenroth, D. Fox and C. E. Rasmussen *Gaussian Processes for Data-Efficient Learning in Robotics and Control* In IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 2, pp. 408-423, 2015.

A. Pekarovskiy, T. Nierhoff, S. Hirche and M. Buss *Dynamically Consistent Online Adaptation of Fast Motions for Robotic Manipulators* In IEEE Transactions on Robotics, vol. 34, no. 1, pp. 166-182, 2018