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

Chattering mitigated sliding mode control of uncertain nonlinear systems

Mark Spiller, Dirk Söffker

Chair of Dynamics and Control, University of Duisburg-Essen , 47057 Duisburg, Germany (mark.spiller@uni-due.de, soeffker@uni-due.de)

Abstract: In this paper chattering mitigated sliding mode control of uncertain nonlinear systems is considered. Concrete knowledge about the system parameters or uncertainty bounds is assumed to be unavailable. A combined sliding mode and data-driven model-free predictive control strategy is proposed. Calculation of the predictive control input is based on a linearized system description. The parameters of the linearized model are estimated online using a Kalman filter and input-output data. The sliding mode controller guarantees boundedness of the tracking error. The switching gain adapts online which avoids the uncertainty bounds of the system to be known. Overestimation of the bounds is avoided by the use of the predictive controller, leading to mitigation of the chattering effect. The effectiveness of the proposed strategy is confirmed by a simulation example.

Keywords: Kalman filters, Sliding mode control, Predictive control, System identification, Adaptation

1. INTRODUCTION

Sliding mode control (SMC) is well established in the field of robust control. It guarantees convergence although modeling inaccuracies with known bounds may be present. Conventional SMC is based on a first order realtionship between the sliding variable and the input. Using a discontinuous control law the sliding variable can be driven to zero, and the state trajectory approaches the sliding surface. Dependent on the definition of the surface the states or tracking error of the system converge to zero (Slotine and Li (1991)). The disadvantage of conventional SMC is the generation of high frequently switching inputs (chattering) which may lead to infeasibility. A boundary layer has been introduced in Slotine (1984) to establish a compromise between precision and chattering attenuation. Higher order SMCs (HSMC), which drive the sliding variable and its derivatives of corresponding order to zero, can improve precision in case of digital control and achieve chattering mitigation. If the order of the SMC is higher than the relative degree of the system the discontinuity in the control law can be avoided (Levant (2003)). However the HSMC approaches require the derivatives of the sliding variables to be known. The derivatives may be estimated by a sliding differentiator as suggest in Levant (2003), but the accuracy of the estimations is affected by measurement noise. Additionally, the HSMC approach of Levant (2003) is restricted to the SISO case, and requires tuning of a gain parameter. Another option to reduce chattering is the usage of exponential power reaching laws. These reaching laws reduce the switching gain in the near of the sliding surface leading to mitigation of the chattering effect. The idea has been introduced in Gao and Hung (1993) and improved regarding reaching time in Fallaha et al. (2010). As the swichting gain depends on the bounds of the uncertainties the chattering effects can increase if the bounds are overestimated. Therefore adaptive SMC approaches (Huang et al. (2008), Plestan et al. (2010), Edwards and Shtessel (2016)) were proposed capable to estimate the uncertainty bounds online.

As SMC guarantees robustness and predictive control is known to be efficient and chattering free, combinations of both approaches have been considered in previous studies. In Garcia-Gabin et al. (2009) a minimization problem based on the sliding variable and its prediction over several time steps is formulated. The minimizing solution of the problem is considered to be the predictive controller which keeps the states on the sliding surface. In order to guarantee that the sliding surface will be reached a conventional SMC with saturation function is added. The approach is applicable to SISO systems. Approaches that use the predictive controller in the reaching phase are presented in Xiao et al. (2007), Xu and Li (2011), and Xu (2015). These approaches avoid the chattering problem as the reaching is achieved based on the optimal predictive control instead of a discontinuous swichting control. Robustness with respect to disturbances can be achieved if the rate of change of the disturbance is known. However the approaches are only applicable to linear systems and only additive disturbances are considered. In Rubagotti et al. (2010) predictive control and SMC are combined based on the framework of integral SMC (I-SMC). Using I-SMC matched uncertainties can be eliminated which simplifies the design of the predictive controller in case of nonlinear systems.

Note that the aforementioned combinations of sliding mode and predictive control assume the system model to be known and do not considered the switching gain to be adaptive which means that knowledge about the uncertainty bounds is required.

In this paper robust control of nonlinear systems with slow dynamics is considered. A combined sliding mode and predictive control approach is proposed. Robustness is achieved through SMC, whereas the predictive controller mitigates the chattering effect and minimizes the input energy. The main advantage of the proposed method is that concrete knowledge about the parameters or the uncertainty bounds is not required. The predictive control input is calculated based on a linear model, which is trained and updated online using a Kalman filter. The combined sliding mode and predictive control approach uses adaption techniques which avoids the uncertainty bounds to be known.

The paper is organized as follows. In Section 2 requirements and the considered class of nonlinear systems are described. System identification based on Kalman filtering is explained in Section 3. Design of the optimal predictive controller is considered in Section 4. The combined sliding mode and predictive control approach is described in Section 5. A numerical example is considered in Section 6.

2. REQUIREMENTS

A nonlinear system

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{u}(t), t), \qquad \mathbf{y}(t) = h(\mathbf{y}(t), t), \tag{1}$$

with states $\mathbf{x}(t) \in \mathbb{R}^l$, measurements $\mathbf{y}(t) \in \mathbb{R}^r$, and inputs $\mathbf{u}(t) \in \mathbb{R}^m$ is considered. The control variable $\mathbf{z}(t) \in \mathbb{R}^m$, and the tracking error $\mathbf{e}(t) \in \mathbb{R}^m$ are defined as

$$\mathbf{z}(t) = \mathbf{L}\mathbf{y}(t), \qquad \mathbf{e}(t) = \mathbf{z}^r(t) - \mathbf{z}(t). \tag{2}$$

where $\mathbf{z}^r(t)$ denotes the reference value. Three assumptions about the behavior of (1) have to be made. First, it is assumed that the discrete-time input-output behavior of (1) can be described by the NARX model

$$\mathbf{y}_{k+1} = q_k(\mathbf{s}_k),$$
$$\mathbf{s}_k = \left[\mathbf{y}_k^T \; \dots \; \mathbf{y}_{k-n_y+1}^T \; \mathbf{u}_k^T \; \dots \; \mathbf{u}_{k-n_u+1}^T\right]^T, \tag{3}$$

with $\mathbf{s}_k \in \mathbb{R}^{(rn_y+mn_u)\times 1}$. Second, the Taylor series expansion of the i-th component of $\mathbf{y}_{k+1}$ given as

$$y_{k+1}^{(i)} = y_k^{(i)} + (\mathbf{s}_k - \mathbf{s}_{k-1})^T \mathbf{D} q_k^{(i)}(\mathbf{s}_{k-1})$$
$$+ \frac{1}{2}(\mathbf{s}_k - \mathbf{s}_{k-1})^T \mathbf{D}^2 q_k^{(i)}(\mathbf{s}_{k-1})(\mathbf{s}_k - \mathbf{s}_{k-1}) + \dots, \tag{4}$$

is considered. The gradient and Hessian of the i-th component of $q_k(\bullet)$ are denoted as $\mathbf{D}q_k^{(i)}$ and $\mathbf{D}^2 q_k^{(i)}$. It is assumed that (1) has slow dynamics and $q_k(\bullet)$ can be approximated well around $\mathbf{y}_k$ by considering only linear terms in (4). Third, consider a sliding variable $\sigma(t) \in \mathbb{R}^m$ to be defined by

$$\sigma^{(i)}(t) = a_{n_\sigma}^{(i)} \frac{\partial^{n_\sigma^{(i)}}}{(\partial t)^{n_\sigma^{(i)}}} e^{(i)}(t)$$
$$+ a_{n_\sigma -1}^{(i)} \frac{\partial^{n_\sigma^{(i)}-1}}{(\partial t)^{n_\sigma^{(i)}-1}} e^{(i)}(t) + \dots + a_0^{(i)} e^{(i)}(t), \tag{5}$$

for $i = 1 \dots m$. System (5) is designed so that it is BIBO stable related to the input $\sigma(t)$ and the output $\mathbf{e}(t)$. The dynamics of the sliding variable resulting from (1), and (5) are assumed to be of the form

$$\dot{\sigma}(t) = \mathbf{v}\left(\mathbf{x}(t), t\right) + \mathbf{W}\left(\mathbf{x}(t), t\right)\mathbf{u}(t), \tag{6}$$

where $\mathbf{W}\left(\mathbf{x}(t), t\right)$ is a diagonal matrix

$$\mathbf{W}\left(\mathbf{x}(t), t\right) = \begin{bmatrix} w^{(1)}(\bullet) & 0 & \dots & 0 \\ 0 & w^{(2)}(\bullet) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w^{(m)}(\bullet) \end{bmatrix}, \tag{7}$$

with known $\mathrm{sign}(w^{(i)}(\bullet))$. It is assumed that finite but unknown uncertainty bounds

$$|v^{(i)}(\bullet)| \le v_M^{(i)}, \qquad 0 < w_m^{(i)} \le |w^{(i)}(\bullet)|, \tag{8}$$

exist.

3. SYSTEM IDENTIFICATION BASED ON ADAPTIVE LINEAR NETWORKS

Considering only the linear parts in (4) a linearization of (3) is obtained as

$$\mathbf{y}_{k+1} \approx \mathbf{A}_k^{(1)}\mathbf{y}_k + \dots + \mathbf{A}_k^{(n_y+1)}\mathbf{y}_{k-n_y} + \mathbf{N}_k\mathbf{u}_k$$
$$+ \mathbf{B}_k^{(1)}\mathbf{u}_{k-1} + \dots + \mathbf{B}_k^{(n_u)}\mathbf{u}_{k-n_u}. \tag{9}$$

The matrices $\mathbf{A}_k^{(i)}, \mathbf{B}_k^{(j)}, \mathbf{N}_k$ in (9) define the transfer function matrix of a linear MIMO system (Isermann and Münchhof (2010)). Based on input-output data the linear system of (9) can be identified by means of e. g. regression (Stenman (1999)), subspace identification (Favoreel et al. (1999)), neural networks (Prasad et al. (1998)). In this work it is suggested to use a linear neural network. The linear system (9) can be rewritten as a neural network

$$\mathbf{y}_{k+1} \approx \mathbf{A}_k\bar{\mathbf{y}}_k + \mathbf{N}_k\mathbf{u}_k + \mathbf{B}_k\bar{\mathbf{u}}_{k-1} + \mathbf{b}_k,$$
$$= \mathbf{Z}_k\mathbf{p}_k = \left[\mathbf{A}_k \; \mathbf{N}_k \; \mathbf{B}_k \; \mathbf{b}_k\right]\left[\bar{\mathbf{y}}_k^T \; \mathbf{u}_k^T \; \bar{\mathbf{u}}_{k-1}^T \; 1\right]^T, \tag{10}$$

with inputs

$$\bar{\mathbf{y}}_k = \left[\mathbf{y}_k^T \; \dots \; \mathbf{y}_{k-n_y}^T\right]^T, \quad \bar{\mathbf{u}}_{k-1} = \left[\mathbf{u}_{k-1}^T \; \dots \; \mathbf{u}_{k-n_u}^T\right]^T,$$

and $\mathbf{u}_k$, weighting matrices $\mathbf{A}_k, \mathbf{B}_k$, and $\mathbf{N}_k$, and bias vector $\mathbf{b}_k$. Input vector $\mathbf{p}_k \in \mathbb{R}^n$ is of dimension $n = r(n_y + 1) + m(n_u + 1) + 1$. The considered network is adaptive as the parameters

$$\zeta_k = \mathrm{vec}(\mathbf{Z}_k), \tag{11}$$

can be estimated and adapted online by means of a Kalman filter so that an updated approximation of the nonlinear system (3) for time step k is available. A well-known Kalman filter based estimation of the network parameters

$$\hat{\zeta}_{k|k} = \mathrm{vec}(\hat{\mathbf{Z}}_{k|k}), \tag{12}$$

is given as (Singhal and Wu (1989))

$$\hat{\zeta}_{k+1|k} = \hat{\zeta}_{k|k}, \tag{13}$$
$$\mathbf{P}_{k+1|k} = \mathbf{P}_{k|k} + \mathbf{Q}, \tag{14}$$
$$\hat{\zeta}_{k+1|k+1} = \hat{\zeta}_{k+1|k} + \mathbf{K}_{k+1}(\mathbf{y}_{k+1} - \mathbf{H}_{k+1}\hat{\zeta}_{k+1|k}), \tag{15}$$
$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^T(\mathbf{H}_{k+1}\mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^T + \mathbf{R})^{-1}, \tag{16}$$
$$\mathbf{P}_{k+1|k+1} = \mathbf{K}_{k+1}\mathbf{R}\mathbf{K}_{k+1}^T + \tag{17}$$
$$(\mathbf{I}_{nr} - \mathbf{K}_{k+1}\mathbf{H}_{k+1})\mathbf{P}_{k+1|k}(\mathbf{I}_{nr} - \mathbf{K}_{k+1}\mathbf{H}_{k+1})^T. \tag{18}$$

The output matrix

$$\mathbf{H}_{k+1} = \mathbf{p}_k^T \otimes \mathbf{I}_r, \tag{19}$$

is obtained by applying the vector operator on (10). The input-output data is assumed to be noise-free

$$\begin{bmatrix}\bar{\mathbf{y}}_k\\ \bar{\mathbf{u}}_{k-1}\end{bmatrix} = \begin{bmatrix}\bar{\mathbf{y}}_k\\ \bar{\mathbf{u}}_{k-1}\end{bmatrix} + \mathbf{r}_k, \tag{20}$$

with $\mathbf{r}_k = 0$, as $\mathbf{r}_k \neq 0$ would affect the output matrix in (19). As Kalman filtering is related to weighted least-squares estimation (Sorenson (1970)) algorithm (13-18) minimizes

$$a = \underset{(\zeta_i^\star)_{i=0}^k}{\arg\min} \|\zeta_0 - \zeta_0^\star\|^2_{\mathbf{P}_0^{-1}} + \sum_{i=0}^{k} \|\mathbf{y}_k - \mathbf{H}_k\zeta_k^\star\|^2_{\mathbf{R}^{-1}} + \sum_{i=0}^{k-1} \|\zeta_{k+1}^\star - \zeta_k^\star\|^2_{\mathbf{Q}^{-1}}, \tag{21}$$

where $a = (\hat{\zeta}_{i|i})_{i=0}^k$ are the Kalman filter estimations. The weighting matrices $\mathbf{Q} = \alpha\mathbf{I}_{nr}$, $\mathbf{R} = \beta\mathbf{I}_r, \alpha \geq 0, \beta > 0$ are design values. Matrix $\mathbf{R}$ determines how exact the estimated network parameters should be fitted to the input-output data, and $\mathbf{Q}$ influences the learning rate of the network. Based on the estimated network parameters a linear state space realization is obtained as

$$\underbrace{\begin{bmatrix}\bar{\mathbf{y}}_{k+1}\\ \bar{\mathbf{u}}_k\\ \hat{\mathbf{b}}_{k+1}\end{bmatrix}}_{\bar{\mathbf{x}}_{k+1}} = \underbrace{\begin{bmatrix}\bar{\mathbf{A}}_{11} & \bar{\mathbf{A}}_{12} & \bar{\mathbf{A}}_{13}\\ 0 & \bar{\mathbf{A}}_{22} & 0\\ 0 & 0 & \mathbf{I}_r\end{bmatrix}}_{\bar{\mathbf{A}}} \underbrace{\begin{bmatrix}\bar{\mathbf{y}}_k\\ \bar{\mathbf{u}}_{k-1}\\ \hat{\mathbf{b}}_k\end{bmatrix}}_{\bar{\mathbf{x}}_k} + \underbrace{\begin{bmatrix}\bar{\mathbf{N}}_1\\ \bar{\mathbf{N}}_2\\ 0\end{bmatrix}}_{\bar{\mathbf{N}}} \mathbf{u}_k,$$
$$\hat{\mathbf{y}}_{k+1} \approx \begin{bmatrix}\mathbf{I}_r & 0 & \dots & 0\end{bmatrix} \bar{\mathbf{y}}_{k+1}, \tag{22}$$

with

$$\bar{\mathbf{A}}_{11} = \begin{bmatrix}\hat{\mathbf{A}}_k\\ \mathbf{T}\end{bmatrix}, \quad \bar{\mathbf{A}}_{12} = \begin{bmatrix}\hat{\mathbf{B}}_k\\ 0\end{bmatrix}, \quad \bar{\mathbf{A}}_{13} = \begin{bmatrix}\mathbf{I}_r\\ 0\end{bmatrix},$$
$$\bar{\mathbf{N}}_1 = \begin{bmatrix}\hat{\mathbf{N}}_k\\ 0\end{bmatrix}, \quad \bar{\mathbf{A}}_{22} = \begin{bmatrix}0\\ \mathbf{S}\end{bmatrix}, \quad \bar{\mathbf{N}}_2 = \begin{bmatrix}\mathbf{I}_m\\ 0\end{bmatrix},$$
$$\mathbf{T} = \begin{bmatrix}\mathbf{I}_r & 0 & \dots & 0 & 0\\ 0 & \mathbf{I}_r & \dots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \dots & \mathbf{I}_r & 0\end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix}\mathbf{I}_m & 0 & \dots & 0 & 0\\ 0 & \mathbf{I}_m & \dots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \dots & \mathbf{I}_m & 0\end{bmatrix},$$

where $\mathbf{T}$ is a $rn_y \times r(n_y+1)$ matrix, and $\mathbf{S}$ is a $m(n_u - 1) \times mn_u$ matrix.

4. MODEL-FREE PREDICTIVE CONTROL

In the following predictive control i. e. minimization of

$$\underset{\vec{\mathbf{u}}}{\arg\min}\, \frac{1}{2}\left(\sum_{i=k}^{k+n_p-1} \mathbf{e}_i^T\mathbf{Q}_i^{PC}\mathbf{e}_i + \sum_{i=k}^{k+n_c-1} \mathbf{u}_i^T\mathbf{R}_i^{PC}\mathbf{u}_i\right),$$
$$\text{s.t}\quad \mathbf{A}_c\vec{\mathbf{u}} \leq \mathbf{b}_c, \quad \vec{\mathbf{u}} = \begin{bmatrix}\mathbf{u}_k & \dots & \mathbf{u}_{k+n_c-2}\end{bmatrix}^T, \tag{23}$$

with tracking error $\mathbf{e}_k$, symmetric weighting matrices $\mathbf{Q}^{PC} \geq 0, \mathbf{R}^{PC} > 0$, prediction horizon n_p, control horizon n_c, with $n_p > n_c$, and constraints $(\mathbf{A}_c, \mathbf{b}_c)$, is considered. As the linear state space description (22) is available the problem reduces to the well-known linear state space based model predictive control (MPC) problem. A brief solution of this problem is given as follows based on Mikuláš (2013), Wang (2009). The state space model (22) is augmented by the reference value leading to

$$\underbrace{\begin{bmatrix}\bar{\mathbf{x}}_{k+1}\\ \mathbf{z}_{k+1}^{ref}\end{bmatrix}}_{\tilde{\mathbf{x}}_{k+1}} = \underbrace{\begin{bmatrix}\bar{\mathbf{A}} & 0\\ 0 & \mathbf{I}_l\end{bmatrix}}_{\tilde{\mathbf{A}}} \underbrace{\begin{bmatrix}\bar{\mathbf{x}}_k\\ \mathbf{z}_k^r\end{bmatrix}}_{\tilde{\mathbf{x}}_k} + \underbrace{\begin{bmatrix}\bar{\mathbf{N}}\\ 0\end{bmatrix}}_{\tilde{\mathbf{N}}} \mathbf{u}_k,$$
$$\mathbf{e}_k = \mathbf{C}\tilde{\mathbf{x}}_k = \begin{bmatrix}\mathbf{L} & 0 & \dots & 0 & -\mathbf{I}_l\end{bmatrix}\tilde{\mathbf{x}}_k. \tag{24}$$

Based on (24) the prediction equation of the tracking error

$$\underbrace{\begin{bmatrix}\mathbf{e}_k\\ \mathbf{e}_{k+1}\\ \vdots\\ \mathbf{e}_{k+n_p-1}\end{bmatrix}}_{\vec{\mathbf{e}}} = \underbrace{\begin{bmatrix}\mathbf{C}\\ \mathbf{C}\tilde{\mathbf{A}}\\ \vdots\\ \mathbf{C}\tilde{\mathbf{A}}^{n_p-1}\end{bmatrix}}_{\boldsymbol{\Psi}} \tilde{\mathbf{x}}_k + \underbrace{\begin{bmatrix}0 & 0 & \dots\\ \mathbf{C}\tilde{\mathbf{N}} & 0 & \dots\\ \vdots & \vdots & \ddots\\ \mathbf{C}\tilde{\mathbf{A}}^{n_p-2}\tilde{\mathbf{N}} & \mathbf{C}\tilde{\mathbf{A}}^{n_p-3}\tilde{\mathbf{N}} & \dots\end{bmatrix}}_{\boldsymbol{\Omega}} \underbrace{\begin{bmatrix}\mathbf{u}_k\\ \vdots\\ \mathbf{u}_{k+n_p-2}\end{bmatrix}}_{\vec{\mathbf{u}}^{,\mathbf{n_p}}}, \tag{25}$$

is obtained. Using (25) problem (23) can be written as a quadratic programm

$$\vec{\mathbf{u}}^\star = \underset{\vec{\mathbf{u}}}{\arg\min}\, \frac{1}{2}\vec{\mathbf{u}}^T\mathbf{G}\vec{\mathbf{u}} + \mathbf{f}^T\vec{\mathbf{u}}, \quad \text{s.t.}\quad \mathbf{A}_c\vec{\mathbf{u}} \leq \mathbf{b}_c, \tag{26}$$

with

$$\mathbf{G} = \mathbf{M}^T\boldsymbol{\Omega}^T\tilde{\mathbf{Q}}^{PC}\boldsymbol{\Omega}\mathbf{M} + \tilde{\mathbf{R}}^{PC}, \quad \mathbf{f}^T = \tilde{\mathbf{x}}_k^T\boldsymbol{\Psi}^T\tilde{\mathbf{Q}}\mathbf{H}\mathbf{M},$$
$$\tilde{\mathbf{Q}}^{PC} = \mathbf{I}_{n_p} \otimes \mathbf{Q}^{PC}, \quad \tilde{\mathbf{R}}^{PC} = \mathbf{I}_{n_c} \otimes \mathbf{R}^{PC}, \quad \vec{\mathbf{u}}^{,\mathbf{n_p}} = \mathbf{M}\vec{\mathbf{u}},$$

where $\mathbf{M}$ is the move blocking matrix keeping $\mathbf{u}_{k+k^*} = \mathbf{u}_{k+n_c}$ fixed for all predictions $k^* > n_c$. Based on $\mathbf{Q}^{PC} \geq 0$, $\mathbf{R}^{PC} > 0$, leading to $\tilde{\mathbf{Q}}^{PC} \geq 0$, $\tilde{\mathbf{R}}^{PC} > 0$, it follows $\mathbf{G} > 0$, so problem (26) is convex (Nocedal and Wright (2006)). Finally, the optimal predictive control input of the next time step is the first entry in $\vec{\mathbf{u}}^\star$ which is denoted as $\mathbf{u}_k^\star$.

5. CHATTERING MITIGATED SLIDING MODE CONTROL

The SMC controller is designed in continuous time space. Consequently, the optimal predicitive control input $\mathbf{u}_k^\star$ is transformed into a time-continuous signal

$$\mathbf{u}^\star(t) = \mathbf{u}_k^\star, \qquad kT_s \leq t < (k+1)T_s, \tag{27}$$

where T_s denotes the sample time. The proposed combined control input $u^{(i)}(t)$ for $i = 1 \dots m$ is

$$u^{(i)}(t) = \eta^{(i)}\left(\sigma^{(i)}(t)\right)u^{\star^{(i)}}(t) - \frac{\mu^{(i)}\left(\sigma^{(i)}(t)\right)}{\text{sign}\left(w^{(i)}\left(\mathbf{x}(t), t\right)\right)}\text{sign}\left(\sigma^{(i)}(t)\right), \tag{28}$$

with

$$\eta^{(i)}\left(\sigma^{(i)}(t)\right) > 0, \qquad \mu^{(i)}\left(\sigma^{(i)}(t)\right) > 0, \tag{29}$$

where $\eta^{(i)}(\bullet)$, and $\mu^{(i)}(\bullet)$ are weighting functions chosen as

- If $|\sigma^{(i)}| > t_\sigma^{(i)}$

$$\eta^{(i)}(\sigma^{(i)}) = 0, \qquad \mu^{(i)}(\sigma^{(i)}) = k_{\text{SMC}}^{(i)}, \tag{30}$$

with

$$\dot{k}_{\text{SMC}}^{(i)}(t) = k_1^{(i)}|\sigma^{(i)}(t)|, \qquad k_1^{(i)} > 0, \tag{31}$$

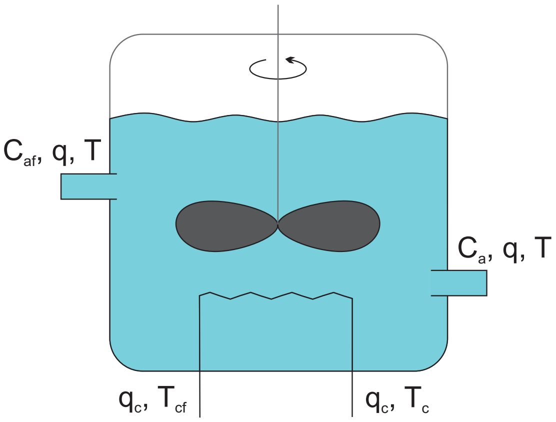

Fig. 1. Continuous stirred tank reactor (CSTR)

- If $|\sigma^{(i)}| \le t_\sigma^{(i)}$

$$\eta^{(i)}(\sigma^{(i)}) = \tilde{\eta}^{(i)}(|\sigma^{(i)}|), \quad \mu^{(i)}(\sigma^{(i)}) = \tilde{\mu}^{(i)}(|\sigma^{(i)}|),$$
$$k_{\mathrm{SMC}}^{(i)} = \tilde{\mu}^{(i)}(t_\sigma^{(i)}). \tag{32}$$

The weighting functions $\tilde{\eta}^{(i)}(\bullet)$, $\tilde{\mu}^{(i)}(\bullet)$ are selected so that the properties

$$\lim_{|\sigma^{(i)}|\to 0} \tilde{\eta}^{(i)}(|\sigma^{(i)}|) = 1, \qquad \frac{d\tilde{\eta}^{(i)}(|\sigma^{(i)}|)}{d|\sigma^{(i)}|} < 0, \tag{33}$$

$$\lim_{|\sigma^{(i)}|\to 0} \tilde{\mu}^{(i)}(|\sigma^{(i)}|) = 0, \qquad \frac{d\tilde{\mu}^{(i)}(|\sigma^{(i)}|)}{d|\sigma^{(i)}|} > 0, \tag{34}$$

hold. Consequently the control is dominated by the MPC solution in the near of the sliding surface, whereas the switching gain increases if it is required to drive the states back to the surface. If the sliding variable exceeds the threshold $t_\sigma^{(i)}$ only the SMC is used for control and the switching gain is adapted based on (31). The adaption of the switching gain follows the adaption law proposed in Huang et al. (2008) which guarantees finite-time stability of $|\sigma^{(i)}|$ with respect to the domain $t_\sigma^{(i)}$ (proven in e. g. Plestan et al. (2010) Corollary 1). Additional to the adaption law of Huang et al. (2008) the resetting (32) is considered to allow the switching gain to decrease to the fixed value $\tilde{\mu}^{(i)}(|\sigma^{(i)}|)$ with $|\sigma^{(i)}| = t_\sigma^{(i)}$.

6. EXAMPLE

The performance of the proposed controller is evaluated based on a simulation example. A chemical reaction of a species A in a continuous stirred tank reactor (Fig. 1) is considered. As shown in Seborg et al. (2010), Magni et al. (2001) the dynamics of the effluent flow concentration $C_A(t) = x_1(t)$ and the reactor temperature $T(t) = x_2(t)$ can be described by

$$\underbrace{\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix}}_{\dot{\mathbf{x}}(t)} = \underbrace{\begin{bmatrix}a(t)\\ b(t)+c(t)\end{bmatrix}}_{d(\mathbf{x}(t))} + \underbrace{\begin{bmatrix}0\\ \frac{UA}{V\rho C_p}\end{bmatrix}}_{g(\mathbf{x}(t))} u(t), \tag{35}$$

$$a(t) = \frac{q}{V}(C_{Af} - x_1(t)) - k_0 x_1(t)\exp\left(-\frac{E}{Rx_2(t)}\right),$$

$$b(t) = \frac{q}{V}(T_f - x_2(t)) + \frac{-\Delta H k_0 x_1(t)}{\rho C_p}\exp\left(-\frac{E}{Rx_2(t)}\right),$$

$$c(t) = \frac{UA}{V\rho C_p}(T_c^{eq} - x_2(t)),$$

with measurements $y(t)$, and control variable $z(t)$ defined as

$$\mathbf{y}(t) = h(\mathbf{x}(t)) = \begin{bmatrix}x_1(t)\\ x_2(t),\end{bmatrix}, \quad z(t) = \kappa(\mathbf{x}(t)) = x_1(t).$$

The parameters of the system are shown in Table 1. The input of the system is the change of the coolant stream temperature $u(t) = \Delta T_c$ related to the nominal value T_c^{eq}. The input saturation is $|u(t)| \le 50\,\mathrm{K} = u_{\max}$. The system is known to have slow dynamics (Magni et al. (2001)). As all states are measured, and the first order derivatives correspond to shifts of one time step in the discrete-time domain, equation (35) can be transformed into the NARX model (3). The relative degree and the Lie derivatives of the system are determined to show that the dynamics of the sliding variable can have the form (6). Based on the Lie derivatives

$$L_d^2\kappa(\mathbf{x}(t)) = \frac{q^2}{V^2}(C_{Af} - x_1(t))$$
$$+ x_1(t)k_0\exp\left(-\frac{E}{Rx_2(t)}\right) \times \left[\frac{q}{V} + k_0\exp\left(-\frac{E}{Rx_2(t)}\right)\right.$$
$$-\frac{q}{Vx_1(t)}(C_{Af} - x_1(t)) - \frac{Eq}{Rx_2(t)^2V}(T_f - x_2(t))$$
$$-x_1(t)\frac{-\Delta H E k_0}{\rho C_p R x_2(t)^2}\exp\left(-\frac{E}{Rx_2(t)}\right)$$
$$\left.-\frac{UAE}{V\rho C_p R x_2(t)^2}(T_c^{eq} - x_2(t))\right],$$

$$L_g L_d\kappa(\mathbf{x}(t)) = -k_0 x_1(t)\frac{UAE}{V\rho C_p R x_2(t)^2}\exp\left(-\frac{E}{Rx_2(t)}\right),$$

the input-output behavior

$$\ddot{z}(t) = L_d^2\kappa(\mathbf{x}(t)) + L_g L_d\kappa(\mathbf{x}(t))\,u(t), \tag{36}$$

is achieved. The relative degree of the system is two. Consider the sliding variable to be defined as

$$\sigma(t) = \dot{e}(t) + \lambda e(t), \qquad \lambda > 0, \tag{37}$$

then the dynamics of the sliding variable are

$$\dot{\sigma}(t) = \ddot{e}(t) + \lambda\dot{e}(t) = v(\mathbf{x}(t), t) + w(\mathbf{x}(t), t)\,u(t), \tag{38}$$

with

$$v(\mathbf{x}(t), t) = \ddot{z}^r(t) - L_d^2\kappa(\mathbf{x}(t)) + \lambda\dot{z}^r(t)$$
$$-\lambda\frac{q}{V}(C_{Af} - x_1(t)) + \lambda k_0 x_1(t)\exp\left(-\frac{E}{Rx_2(t)}\right),$$
$$w(\mathbf{x}(t), t) = -L_g L_d\kappa(\mathbf{x}(t)).$$

In order to achieve the control law $\operatorname{sign}(-L_g L_d\kappa(\bullet))$ must be determined. Considering $L_g L_d\kappa(\bullet)$ and interpreting the physical meaning of the parameters based on Table 1 it turns out that $\operatorname{sign}(-L_g L_d\kappa(\bullet))$ equals one. This result can be obtained without knowing concrete values of the parameters. Alternatively, it would also be possible to try to use $\operatorname{sign}(-L_g L_d\kappa(\bullet)) = -1$ or $\operatorname{sign}(-L_g L_d\kappa(\bullet)) = 1$ in combination with the controller. One option will work for sure as $x_1(t) > 0$ holds true, and $\operatorname{sign}(-L_g L_d\kappa(\bullet))$ can not change dependent on the states. Finally, the control law can be stated as

$$u(t) = \eta(\sigma(t))\,u^\star(t) - \mu(\sigma(t))\operatorname{sign}(\sigma(t)), \tag{39}$$

Table 1. CSTR process parameters (Magni et al. (2001))

Parameter	Symbol	Value
Tank volume	V	100 l
Feed flow rate	q	100 l/min
Feed concentration	C_{Af}	1 mol/l
Feed temperature	T_f	350 K
Density	ρ	1000 g/l
Enthalpy	$-\Delta H$	5×10^4 J/mol
Exponential factor	$\frac{E}{R}$	8750 K
Frequency factor	k_0	7.2×10^{10} min^{-1}
Heat transfer characteristic	UA	5×10^4 J/minK
Specific heat	C_p	0.239 J/gK
Coolant flow temperature	T_c^{eq}	300 K

where the weighting functions

- If $|\sigma(t)| > 1$ then

$$\eta\left(\sigma(t)\right) = 0, \quad \mu\left(\sigma(t)\right) = k_{\text{SMC}}, \tag{40}$$

$$\dot{k}_{\text{SMC}}(t) = 10|\sigma(t)|\,, \tag{41}$$

- If $|\sigma(t)| \leq 1$ then

$$\eta\left(\sigma(t)\right) = \frac{1}{|\sigma(t)| + 1}, \tag{42}$$

$$\mu\left(\sigma(t)\right) = 10|\sigma(t)|\,, \qquad k_{\text{SMC}}(t) = 10, \tag{43}$$

with $k_{\text{SMC}}(t_0) = 20$ have been determined by trial and error. Process (35) is discretized based on Euler method using a sampling time of 1 s. The simulation has a duration of $T_{\text{sim}} = 15$ min. For the initialization of the states $x_1(t_0) = 0.875$ mol/l, and $x_2(t_0) = 325$ K is considered. The number of delayed inputs and outputs in the network are selected as $n_y = 3, n_u = 2$. The network weights are initialized with $\hat{\mathbf{x}}_0 = \mathbf{I}_{nr\times 1}$, $\mathbf{P}_{0|0} = \mathbf{I}_{nr\times nr} \times 10^{10}$. The learning rate is considered to be $\alpha = 0.01$, and β is chosen as $\beta = 0.001$. The network is initially trained based on the system outputs generated by $u(t^\star) = A_t(t^\star) \times \cos\left(\omega_t(t^\star)t^\star\right), t^\star = 0...20$ min, where $\omega_t(t^\star)$ varies between $\frac{2\pi}{50}$ and $\frac{2\pi}{3000}$, and $A_t(t^\star)$ varies between 0.5 and 2.5. The weighting matrices considered for the model-free predictive control approach are $\mathbf{Q}^{PC} = 1 \times \mathbf{I}_{l\times l}, \mathbf{R}^{PC} = 0.001 \times \mathbf{I}_{m\times m}$. Matrix $\mathbf{L}$ is $\mathbf{L} = \begin{bmatrix}1 & 0\end{bmatrix}$. The prediction and control horizon are chosen as $n_p = 10, n_c = 9$. For the sliding dynamics $\lambda = 0.05$ is considered. The reference values for tracking are

$$z^r(t) = \begin{cases} 0.8\,\text{mol/l} & \text{if } t \leq 4\,\text{min}, \\ 0.75\,\text{mol/l} & \text{if } 4\,\text{min} < t \leq 8\,\text{min}, \\ 0.7\,\text{mol/l} & \text{if } 8\,\text{min} < t \leq 12\,\text{min}, \\ 0.85\,\text{mol/l} & \text{if } 12\,\text{min} < t \leq 15\,\text{min}. \end{cases} \tag{44}$$

The proposed chattering mitigated SMC (CM-SMC) approach (39) is compared to the adaptive SMC (A-SMC) approach of Plestan et al. (2010), and the pure model-free predictive control (MF-PC) approach (26). According to "algorithm 1" in Plestan et al. (2010) the A-SMC approach is given as

$$u_{\text{A-SVSF}}(t) = -k_{\text{A-SMC}}(t) \times \text{sign}\left(\sigma(t)\right), \tag{45}$$

- If $|\sigma(t)| > 1$ then

Table 2. Performance evaluation

	CM-SMC	A-SMC	MF-PC
$\int(e(t)^2/T_{\text{sim}})dt$	584.7	623.6	606.1
$\int(u(t)^2/T_{\text{sim}})dt$	331.8	742.5	565.0

$$\dot{k}_{\text{A-SMC}}(t) = 10|\sigma(t)|\,, \tag{46}$$

- If $|\sigma(t)| \leq 1$ then

$$k_{\text{A-SMC}}(t) = \bar{k}_{\text{A-SMC}}|\gamma(t)| + 2, \tag{47}$$

$$100\dot{\gamma}(t) + \gamma(t) = \text{sign}\left(\sigma(t)\right), \tag{48}$$

with $k_{\text{A-SMC}}(t_0) = 20, \gamma(t_0) = 0, \bar{k}_{\text{A-SMC}} = k_{\text{A-SMC}}(t^\star)$, where $k_{\text{A-SMC}}(t^\star)$ denotes the last value of $k_{\text{A-SMC}}(t)$ before $|\sigma(t)|$ is switching from $|\sigma(t)| > 1$ to $|\sigma(t)| \leq 1$.

The output trajectories of the controlling approaches are visualized in Fig. 2. Performance values are given in Table 2. The A-SMC approach is slow, shows chattering effects, and is more inaccurate in comparison to CM-SMC. The PC approach is stationary accurate but shows overshooting behavior in case of switching reference. The proposed CF-SMC approach can reject the overshooting and is stationary accurate. The generated inputs of the controllers are visualized in Fig. 3. The A-SMC approach shows strong chattering effects. The proposed CM-SMC approach only shows chattering during the rejection of the overshooting. In Fig. 4 the values of the weighting functions during the rejection of the overshooting are shown. When the reference signal changes the value of the swichting gain increases and forces the states back on the sliding surface. Near the surface the values of the swichting gain are scaled down so that chattering is avoided.

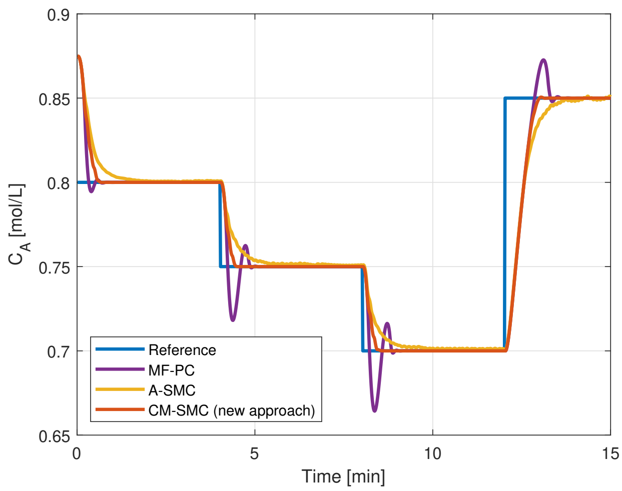

Fig. 2. Output trajectories of the control approaches

7. CONCLUSION

In this paper chattering mitigated sliding mode control of nonlinear systems with slow dynamics is considered. An approach that attenuates the chattering effects without the requirement of knowing concrete parameter values or uncertainty bounds is proposed. In the numerical simulation the suggested approach shows superior performance in comparison to a conventional adaptive sliding mode controller and a pure model-free predictive control approach.

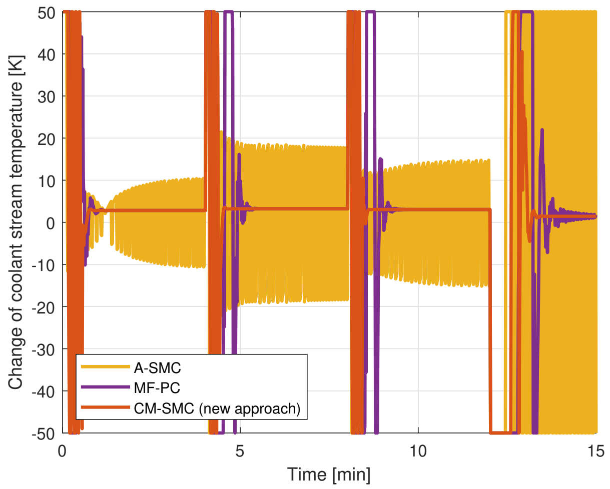

Fig. 3. Inputs generated by the control approaches

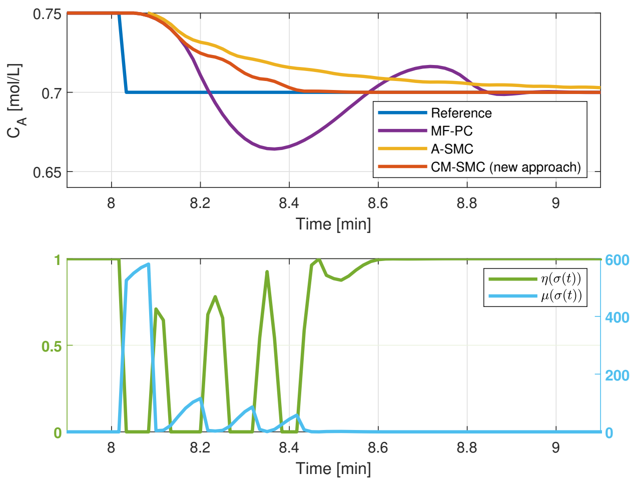

Fig. 4. Rejection of overshooting using switching gain

REFERENCES

Edwards, C. and Shtessel, Y.B. (2016). Adaptive continuous higher order sliding mode control. *Automatica*, 65, 183–190.

Fallaha, C.J., Saad, M., Kanaan, H.Y., and Al-Haddad, K. (2010). Sliding-mode robot control with exponential reaching law. *IEEE Trans. on Industrial Electronics*, 58(2), 600–610.

Favoreel, W., De Moor, B., Gevers, M., and Van Overschee, P. (1999). Closed-loop model-free subspace-based LQG-design. In *Proc. of the 7th IEEE Mediterranean Conference on Control and Automation, June*, 28–30.

Gao, W. and Hung, J.C. (1993). Variable structure control of nonlinear systems: A new approach. *IEEE Trans. on Industrial Electronics*, 40(1), 45–55.

Garcia-Gabin, W., Zambrano, D., and Camacho, E.F. (2009). Sliding mode predictive control of a solar air conditioning plant. *Control Engineering Practice*, 17(6), 652–663.

Huang, Y.J., Kuo, T.C., and Chang, S.H. (2008). Adaptive sliding-mode control for nonlinearsystems with uncertain parameters. *IEEE Trans. on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 38(2), 534–539.

Isermann, R. and Münchhof, M. (2010). *Identification of dynamic systems: An introduction with applications.* Springer Science & Business Media.

Levant, A. (2003). Higher-order sliding modes, differentiation and output-feedback control. *International Journal of Control*, 76(9-10), 924–941.

Magni, L., De Nicolao, G., Magnani, L., and Scattolini, R. (2001). A stabilizing model-based predictive control algorithm for nonlinear systems. *Automatica*, 37(9), 1351–1362.

Mikuláš, O. (2013). Quadratic programming algorithms for fast model-based predictive control. *Bachelor thesis, Czech Technical University in Prague, Prague.*

Nocedal, J. and Wright, S. (2006). *Numerical optimization.* Springer Science & Business Media.

Plestan, F., Shtessel, Y., Bregeault, V., and Poznyak, A. (2010). New methodologies for adaptive sliding mode control. *International Journal of Control*, 83(9), 1907–1919.

Prasad, G., Swidenbank, E., and Hogg, B. (1998). A neural net model-based multivariable long-range predictive control strategy applied in thermal power plant control. *IEEE Trans. on Energy Conversion*, 13(2), 176–182.

Rubagotti, M., Raimondo, D.M., Ferrara, A., and Magni, L. (2010). Robust model predictive control with integral sliding mode in continuous-time sampled-data nonlinear systems. *IEEE Trans. on Automatic Control*, 56(3), 556–570.

Seborg, D.E., Mellichamp, D.A., Edgar, T.F., and Doyle III, F.J. (2010). *Process Dynamics and Control.* John Wiley & Sons.

Singhal, S. and Wu, L. (1989). Training multilayer perceptrons with the extended Kalman algorithm. In *Advances in Neural Information Processing Systems*, 133–140.

Slotine, J.J.E. (1984). Sliding controller design for nonlinear systems. *International Journal of Control*, 40(2), 421–434.

Slotine, J.J.E. and Li, W. (1991). *Applied Nonlinear Control*, volume 199. Prentice hall Englewood Cliffs, N.J.

Sorenson, H.W. (1970). Least-squares estimation: From Gauss to Kalman. *IEEE Spectrum*, 7(7), 63–68.

Stenman, A. (1999). Model-free predictive control. In *Proc. of the 38th IEEE Conference on Decision and Control*, volume 4, 3712–3717. IEEE.

Wang, L. (2009). *Model predictive control system design and implementation using MATLAB®.* Springer Science & Business Media.

Xiao, L., Su, H., and Chu, J. (2007). Sliding mode prediction tracking control design for uncertain systems. *Asian Journal of Control*, 9(3), 317–325.

Xu, Q. (2015). Digital integral terminal sliding mode predictive control of piezoelectric-driven motion system. *IEEE Trans. on Industrial Electronics*, 63(6), 3976–3984.

Xu, Q. and Li, Y. (2011). Model predictive discrete-time sliding mode control of a nanopositioning piezostage without modeling hysteresis. *IEEE Trans. on Control Systems Technology*, 20(4), 983–994.