



On the Relation Between Smooth Variable Structure and Adaptive Kalman Filter

Mark Spiller* and Dirk Söffker

Chair of Dynamics and Control, University of Duisburg-Essen, Duisburg, Germany

This article is addressed to the topic of robust state estimation of uncertain nonlinear systems. In particular, the smooth variable structure filter (SVSF) and its relation to the Kalman filter is studied. An adaptive Kalman filter is obtained from the SVSF approach by replacing the gain of the original filter. Boundedness of the estimation error of the adaptive filter is proven. The SVSF approach and the adaptive Kalman filter achieve improved robustness against model uncertainties if filter parameters are suitably optimized. Therefore, a parameter optimization process is developed and the estimation performance is studied.

Keywords: state estimation, nonlinear systems, stochastic systems, Kalman filtering, control

OPEN ACCESS

Edited by:

Lili Lei,
Nanjing University, China

Reviewed by:

Xanthi Pedeli,
Athens University of Economics and
Business, Greece

Yong Xu,
Northwestern Polytechnical
University, China

*Correspondence:

Mark Spiller
mark.spiller@uni-due.de

Specialty section:

This article was submitted to
Dynamical Systems,
a section of the journal
Frontiers in Applied Mathematics
and Statistics

Received: 20 July 2020

Accepted: 16 October 2020

Published: 27 November 2020

Citation:

Spiller M and Söffker D (2020) On the
Relation Between Smooth Variable
Structure and Adaptive Kalman Filter.
Front. Appl. Math. Stat. 6:585439.
doi: 10.3389/fams.2020.585439

1. INTRODUCTION

State estimation plays an important role in the field of control. System states are required for the calculation of state feedback controllers, exact input/output linearizations, equivalent control, backstepping control, etc. Noise reduction of signals is desirable to improve performance of sliding mode approaches under real conditions. Combined input-state estimation is useful for the estimation and rejection of unknown exogenous inputs. Additionally, robust model-based fault detection and localization approaches can be designed based on filters.

Related to linear systems minimum variance unbiased [1, 2] and augmented state filters [3, 4] can be used for combined input-state estimation. In case of known uncertainty bounds robust Kalman filters [5] can improve state estimation of uncertain linear systems. In the field of H_∞ filtering robustness of the state estimation may be improved by minimizing the effect of worst possible energy-bounded disturbances on the estimation error [6]. Multiple-model approaches [7] are a powerful tool for state estimation of uncertain systems. Combining them with the particle filter allows state estimation of nonlinear systems [8].

The smooth variable structure filter (SVSF) introduced in Ref. 9 is an approach for state estimation of uncertain nonlinear systems. Several applications of SVSF can be found in the literature. The filter has been applied to estimate the states and parameters of an uncertain linear hydraulic system in Ref. 9. A multiple-model approach has been formulated for fault detection e.g., leakage of the hydraulic system [10]. The state of charge and state of health of batteries is estimated in Refs. 11 and 12. A multiple-model approach has been applied for target tracking in Ref. 13 and a SVSF based probabilistic data association (PDA) approach has been proposed for tracking in cluttered environment [14]. For multiple object tracking a SVSF based joint-PDA approach has been developed [15]. Online multiple vehicle tracking on real road scenarios has been investigated in Ref. 16. Several SVSF based simultaneous localization and mapping algorithms have been proposed e.g., Refs. 17–19. Training of neural networks based on SVSF and classification of engine faults has been studied in Ref. 20. Dual estimation of states and model parameters has been

considered in Ref. 21. The estimation strategy works as follows. The bi-section method, the shooting method, and SVSF are combined. The bi-section and shooting method are applied to determine best-fitting model parameter combinations. The obtained model is used by SVSF to estimate the system states. To apply the bi-section method the measurement signals are divided into segments in which the model parameters remain nearly constant. In comparison to the Kalman filter the SVSF approach facilitates detection of these segments based on an evaluation of chattering process.

As described in Ref. 9 the SVSF approach uses a switching gain to drive the estimated state trajectory into a region around the true state trajectory called existence subspace. Due to measurement noise chattering occurs in the existence subspace. By introducing a boundary layer, similar to the saturation function approach of sliding mode control, the high frequency switching in the existence subspace can be attenuated. The claimed advantage of SVSF approach is that if the boundary layer is not used, the filter guarantees to reach the existence subspace for sure although an imprecise system model may be used. It is shown in Ref. 22 that in case of not using the boundary layer the estimations of the filter converge to the measurements, which guarantees the estimation error to be bounded. However, this is a trivial result because if the estimations are equal to the measurements and every state is measured, than the filter is not required. If the boundary layer is used, the estimations diverge from the noisy measurements and estimation performance improves. Unfortunately, it has never been proven that the SVSF with boundary layer has bounded estimation error in case of using an imprecise model.

As already mentioned a serious limitation of the SVSF approach is that all system states have to be measured. Additionally, the measurement model is required to be linear. However, related to tracking a linear measurement model may be achieved by applying a measurement conversion [23] and measurements of the vehicle velocities could also be derived from measured positions.

Another problem of SVSF results from the dependency of the estimation performance on the width of the introduced smoothing boundary layer. In Ref. 24 an estimation error model for the SVSF is proposed and in Ref. 25 the estimation error is minimized according to the smoothing boundary layer width. A maximum a posteriori estimation of the noise statistics of the error model is discussed in Ref. 26. However, the derived estimation error model and the related approaches require the system to be linear and precisely known which contradicts the idea of robustness.

In our previous publication [22] a new tuning parameter for the SVSF approach was introduced to achieve online optimization of the estimation performance. In this paper the relation between the SVSF approach and the Kalman filter is studied. An adaptive Kalman filter is obtained from the SVSF approach by replacing the original filter gain. The estimation performance of SVSF and the adaptive filter variant is compared with one another. Therefore, a parameter optimization scheme is proposed. In the simulation results the adaptive Kalman filter

shows superior performance compared to the original SVSF approach.

The paper is organized as follows. In **Section 2** the preliminaries and the original SVSF approach are discussed. In **Section 3** the relation of SVSF and Kalman filter is studied. Parameter optimization of the Kalman filter leading to an adaptive filtering approach is considered in **Section 4**. The stability of the adaptive filter is studied in **Section 5**. A performance evaluation of SVSF and adaptive Kalman filter is provided in **Section 6**.

Notations. An overview of the notations used within the paper is given in **Table 1**.

2. PROBLEM FORMULATION AND PREVIOUS WORK

Consider the dynamics of a nonlinear system to be exactly described by the discrete-time model

$$\mathbf{x}_{k+1} = f_k(\mathbf{x}_k, \mathbf{u}_k), \quad (1)$$

$$\mathbf{y}_k = \mathbf{x}_k + \mathbf{r}_k, \quad (2)$$

with states $\mathbf{x}_k \in \mathbb{R}^n$, inputs $\mathbf{u}_k \in \mathbb{R}^m$, and outputs $\mathbf{y}_k \in \mathbb{R}^n$. The process $\mathbf{r}_k \in \mathbb{R}^n$ is assumed to be a white noise process with independent samples described by the covariances $E(\mathbf{r}_k \mathbf{r}_k^T) = \mathbf{R} > 0$, and $E(\mathbf{r}_i \mathbf{r}_j^T) = 0$, for $i \neq j$, and the mean $E(\mathbf{r}_k) = 0$. The measurement model Eq. 2 can be obtained from any model of the form $\tilde{\mathbf{y}}_k = \mathbf{H}_k \mathbf{x}_k + \tilde{\mathbf{r}}_k$ as the considered SVSF approach requires \mathbf{H}_k to be invertible [9]. Consider \check{f}_k to be a nominal description of system (1, 2), which may differ from the true behavior. According to Ref. 9 an estimation $\hat{\mathbf{x}}$ of the system states \mathbf{x} can be obtained using the SVSF algorithm

$$\mathbf{e}_{y_{k|k}} = \mathbf{y}_k - \hat{\mathbf{x}}_{k|k}, \quad (3)$$

$$\hat{\mathbf{x}}_{k+1|k} = \check{f}_k(\hat{\mathbf{x}}_{k|k}, \mathbf{u}_k), \quad (4)$$

$$\mathbf{e}_{y_{k+1|k}} = \mathbf{y}_{k+1} - \hat{\mathbf{x}}_{k+1|k}, \quad (5)$$

$$\mathbf{M}_{k+1} = \left(\left| \mathbf{e}_{y_{k+1|k}} \right| + \Phi \left| \mathbf{e}_{y_{k+1|k}} \right| \right) \circ \text{sat}(\mathbf{e}_{y_{k+1|k}}, \Psi), \quad (6)$$

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{M}_{k+1}, \quad (7)$$

where the operator “ \circ ” denotes the Schur product, Φ and Ψ are diagonal matrices with Φ_{ii} , Ψ_{ii} denoting the i th diagonal element, and the i th element of vector $\text{sat}(\mathbf{e}_{y_{k+1|k}}, \Psi)$ is defined as

$$\text{sat}(\mathbf{a}, \mathbf{A})_i = \begin{cases} 1 & \text{for } \frac{\mathbf{a}_i}{\mathbf{A}_{ii}} > 1, \\ -1 & \text{for } \frac{\mathbf{a}_i}{\mathbf{A}_{ii}} < -1, \\ \frac{\mathbf{a}_i}{\mathbf{A}_{ii}} & \text{for } -1 \leq \frac{\mathbf{a}_i}{\mathbf{A}_{ii}} \leq 1, \end{cases} \quad (8)$$

where $i \in \{1, 2, \dots, n\}$. According to Ref. 22 the SVSF algorithm (Eqs 3–7) reduces to

$$\hat{\mathbf{x}}_{k+1|k} = \check{f}_k(\hat{\mathbf{x}}_{k|k}, \mathbf{u}_k), \quad (9)$$

$$\mathbf{e}_{y_{k+1|k}} = \mathbf{y}_{k+1} - \hat{\mathbf{x}}_{k+1|k}, \quad (10)$$

TABLE 1 | Nomenclature.

Symbol	Description
$\hat{\cdot}$	Estimated quantity
$f(\cdot)$	Real system model
$\tilde{f}(\cdot)$	Nominal system model (imprecise)
$E(\cdot)$	Expectation value
$(\cdot)^T$	Transpose operator
$\text{tr}(\mathbf{A})$	Trace of matrix \mathbf{A}
\mathbf{a}_i	Element i of vector \mathbf{a}
$\mathbf{A}_{ii} = (\mathbf{A})_{ii}$	Diagonal element i of square matrix \mathbf{A}
$\mathbf{A} > 0$	Positive definite matrix \mathbf{A}
$\mathbf{A} \geq 0$	Positive semi-definite matrix \mathbf{A}
$ \mathbf{a} $	Absolute value applied element-wise on vector \mathbf{a}
\mathbf{a}	Euclidean vector norm $\sqrt{\mathbf{a}^T \mathbf{a}}$
$\ \mathbf{a}\ _w$	Weighted Euclidean vector norm $\sqrt{\mathbf{a}^T \mathbf{W} \mathbf{a}}$
$\ \mathbf{A}\ $	Frobenius matrix norm $\sqrt{\text{tr}(\mathbf{A}^T \mathbf{A})}$
$\mathbf{a} \circ \mathbf{b}$	Schur product of vectors \mathbf{a} and \mathbf{b}
$\mathbf{A} \otimes \mathbf{B}$	Kronecker product of matrices \mathbf{A} and \mathbf{B}
$\mathbf{a} = \text{vec}(\mathbf{A})$	Vectorization by concatenating matrix columns of \mathbf{A}
$\mathbf{A} = \text{vec}^{-1}(\mathbf{a})$	Inverse vector operator (dimension of \mathbf{A} assumed as known)
$\text{sgn}(\mathbf{a})$	Signum function applied element-wise on vector \mathbf{a}
$\mathbf{A} = \text{diag}(\mathbf{a})$	Build diagonal matrix \mathbf{A} with $\mathbf{A}_{ii} = \mathbf{a}_i$
$\lambda_{\min}\{\mathbf{A}\}$	Minimal eigenvalue of matrix \mathbf{A}

$$\hat{\mathbf{x}}_{k+1|k+1} = \mathbf{y}_{k+1} - \tilde{\mathbf{M}}_{k+1} \mathbf{e}_{y_{k+1|k}}, \tag{11}$$

$$\tilde{\mathbf{M}}_{k+1} = \text{diag}(\mathbf{I}_{n \times 1} - |\text{sat}(\mathbf{e}_{y_{k+1|k}}, \Psi)|), \tag{12}$$

in case of $\Phi = 0_{n \times n}$.

Theorem 1 of Ref. 9 proves that the output estimation error $\mathbf{e}_{y_{k|k}}$ of algorithm (Eqs 3–7) approaches zero if instead of the gain \mathbf{M}_{k+1} the gain $\mathbf{M}_{k+1}^* = (|\mathbf{e}_{y_{k+1|k}}| + \Phi |\mathbf{e}_{y_{k|k}}|) \circ \text{sgn}(\mathbf{e}_{y_{k+1|k}})$ is used. As a consequence the estimations of the filter equal the measurements, i.e., $\mathbf{e}_{y_{k|k}} = \mathbf{y}_k - \hat{\mathbf{x}}_{k|k} = 0$. This is a trivial result as all states are required to be measured. So one might just use the measurements instead of the filter estimates. The estimations of the filter diverge from the measurements and estimation performance improves if a boundary layer is introduced and gain \mathbf{M}_{k+1}^* is replaced by \mathbf{M}_{k+1} . However, it has neither been proven that algorithm (Eqs 3–7) with gain \mathbf{M}_{k+1} has a bounded estimation error, nor it has been shown that the introduced boundary layer minimizes the squared estimation error or some other performance criterion.

3. RELATION BETWEEN SMOOTH VARIABLE STRUCTURE AND KALMAN FILTER

In this section the stochastic gain $\tilde{\mathbf{M}}_{k+1}$ of SVSF approach is replaced by a deterministic but yet undefined gain \mathbf{K}_{k+1} . Using the deterministic gain the estimation error covariance matrix of the filter is determined. By minimizing the mean squared estimation error (MSE) the deterministic gain \mathbf{K}_{k+1} becomes the specific optimal one $\mathbf{K}_{k+1}^{\text{opt}}$. Replacing the original gain $\tilde{\mathbf{M}}_{k+1}$ of SVSF by the optimal one $\mathbf{K}_{k+1}^{\text{opt}}$ gives a direct link to the Kalman filter.

From Eqs 1, 2, and 9–12 it follows that the state estimation error and the output estimation error can be determined as

$$\mathbf{e}_{x_{k+1|k+1}} = \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k+1} = \mathbf{K}_{k+1} \mathbf{e}_{y_{k+1|k}} - \mathbf{r}_{k+1}, \tag{13}$$

and

$$\mathbf{e}_{y_{k+1|k}} = \mathbf{y}_{k+1} - \hat{\mathbf{x}}_{k+1|k} = \mathbf{e}_{x_{k+1|k}} + \mathbf{r}_{k+1}, \tag{14}$$

where \mathbf{K}_{k+1} is a deterministic but yet undefined gain. To derive the optimal gain that minimizes the MSE an expression of the a posteriori error covariance dependent on \mathbf{K}_{k+1} is required to be derived. However, first of all the calculation of the output error covariance is considered. Inserting Eq. 14 into the definition of the output error covariance $\mathbf{S}_{k+1} = E(\mathbf{e}_{y_{k+1|k}} \mathbf{e}_{y_{k+1|k}}^T)$ gives

$$\mathbf{S}_{k+1} = E\left(\left(\mathbf{e}_{x_{k+1|k}} + \mathbf{r}_{k+1}\right)\left(\mathbf{e}_{x_{k+1|k}} + \mathbf{r}_{k+1}\right)^T\right). \tag{15}$$

Expanding Eq. 15 and considering the definition of the a priori estimation error covariance $\mathbf{P}_{k+1|k} = E(\mathbf{e}_{x_{k+1|k}} \mathbf{e}_{x_{k+1|k}}^T)$ and the stationary measurement noise covariance $\mathbf{R} = E(\mathbf{r}_{k+1} \mathbf{r}_{k+1}^T)$ yields

$$\mathbf{S}_{k+1} = \mathbf{P}_{k+1|k} + \mathbf{R} + E\left(\mathbf{r}_{k+1} \mathbf{e}_{x_{k+1|k}}^T\right) + E\left(\mathbf{e}_{x_{k+1|k}} \mathbf{r}_{k+1}^T\right). \tag{16}$$

The value of the remaining expectations in Eq. 16 is studied as follows. First of all the a priori estimation error $\mathbf{e}_{x_{k+1|k}}$ is known to directly depend on the noise realizations \mathbf{r}_j with $j \in \{0, 1, \dots, k\}$ but not on the realization \mathbf{r}_{k+1} . In addition \mathbf{r}_j with $j \in \{0, 1, \dots, k\}$ and \mathbf{r}_{k+1} are independent of each other due to the independent white noise assumption. Finally, \mathbf{r}_{k+1} can not have any effect on $\mathbf{e}_{x_{k+1|k}}$ and both random variables are stochastically independent. It follows $E(\mathbf{r}_{k+1} \mathbf{e}_{x_{k+1|k}}^T) = E(\mathbf{r}_{k+1})E(\mathbf{e}_{x_{k+1|k}}^T)$ and $E(\mathbf{e}_{x_{k+1|k}} \mathbf{r}_{k+1}^T) = E(\mathbf{e}_{x_{k+1|k}})E(\mathbf{r}_{k+1}^T)$. From the zero-mean assumption of the noise i.e., $E(\mathbf{r}_k) = 0$ it follows

$$\mathbf{S}_{k+1} = \mathbf{P}_{k+1|k} + \mathbf{R}. \tag{17}$$

Next the expression for the a posteriori estimation error covariance is considered. Inserting Eq. 13 into the definition of the a posteriori estimation error covariance $\mathbf{P}_{k+1|k+1} = E(\mathbf{e}_{x_{k+1|k+1}} \mathbf{e}_{x_{k+1|k+1}}^T)$ leads to

$$\mathbf{P}_{k+1|k+1} = E\left(\left(\mathbf{K}_{k+1} \mathbf{e}_{y_{k+1|k}} - \mathbf{r}_{k+1}\right)\left(\mathbf{e}_{y_{k+1|k}}^T \mathbf{K}_{k+1}^T - \mathbf{r}_{k+1}^T\right)\right). \tag{18}$$

Expanding Eq. 18 and considering the definition of the output error covariance \mathbf{S}_{k+1} yields

$$\begin{aligned} \mathbf{P}_{k+1|k+1} &= \mathbf{K}_{k+1} \mathbf{S}_{k+1} \mathbf{K}_{k+1}^T + \mathbf{R} - E\left(\mathbf{r}_{k+1} \mathbf{e}_{y_{k+1|k}}^T \mathbf{K}_{k+1}^T\right) \\ &\quad - E\left(\mathbf{K}_{k+1} \mathbf{e}_{y_{k+1|k}} \mathbf{r}_{k+1}^T\right). \end{aligned} \tag{19}$$

Based on Eq. 14 the two remaining expectations in Eq. 19 can be written as

$$\begin{aligned} E\left(\mathbf{r}_{k+1} \left(\mathbf{e}_{x_{k+1|k}}^T + \mathbf{r}_{k+1}^T\right) \mathbf{K}_{k+1}^T\right) &= \mathbf{R} \mathbf{K}_{k+1}^T, \\ E\left(\mathbf{K}_{k+1} \left(\mathbf{e}_{x_{k+1|k}} + \mathbf{r}_{k+1}\right) \mathbf{r}_{k+1}^T\right) &= \mathbf{K}_{k+1} \mathbf{R}, \end{aligned} \tag{20}$$

where $E(\mathbf{r}_{k+1} \mathbf{e}_{x_{k+1|k}}^T)$ and $E(\mathbf{e}_{x_{k+1|k}} \mathbf{r}_{k+1}^T)$ again vanish due to the stochastic independency of \mathbf{r}_{k+1} and $\mathbf{e}_{x_{k+1|k}}$. Finally, the a posteriori error covariance is achieved as

$$\mathbf{P}_{k+1|k+1} = \mathbf{K}_{k+1} \mathbf{S}_{k+1} \mathbf{K}_{k+1}^T + \mathbf{R} - \mathbf{R} \mathbf{K}_{k+1}^T - \mathbf{K}_{k+1} \mathbf{R}. \tag{21}$$

In the following minimization of MSE is considered. Based on Eq. 21 the error covariance can also be written as

$$\mathbf{P}_{k+1|k+1} = \mathbf{K}_{k+1} \mathbf{S}_{k+1} \mathbf{K}_{k+1}^T + \mathbf{R} - \mathbf{R} \mathbf{S}_{k+1}^{-1} \mathbf{S}_{k+1} \mathbf{K}_{k+1}^T - \mathbf{K}_{k+1} \mathbf{S}_{k+1} \mathbf{S}_{k+1}^{-1} \mathbf{R}. \tag{22}$$

By adding $\mathbf{R} \mathbf{S}_{k+1}^{-1} \mathbf{R} - \mathbf{R} \mathbf{S}_{k+1}^{-1} \mathbf{R}$ expression Eq. 22 can be rewritten as

$$\mathbf{P}_{k+1|k+1} = \mathbf{R} - \mathbf{R} \mathbf{S}_{k+1}^{-1} \mathbf{R} + (\mathbf{K}_{k+1} - \mathbf{R} \mathbf{S}_{k+1}^{-1}) \mathbf{S}_{k+1} (\mathbf{K}_{k+1} - \mathbf{R} \mathbf{S}_{k+1}^{-1})^T. \tag{23}$$

The trace of $\mathbf{P}_{k+1|k+1}$ gives the MSE of the a posteriori estimation. The last term in Eq. 23 is at least a positive semidefinite matrix with a trace greater equal zero as $\mathbf{S}_{k+1} > 0$. All other terms do not depend on \mathbf{K}_{k+1} . Consequently, the minimum of MSE is achieved by

$$\mathbf{K}_{k+1}^{\text{opt}} = \mathbf{R} \mathbf{S}_{k+1}^{-1} = \mathbf{R} (\mathbf{P}_{k+1|k} + \mathbf{R})^{-1}, \tag{24}$$

as the trace of the last term is zero for $\mathbf{K}_{k+1}^{\text{opt}}$. The connection between algorithm (Eqs 3–7) and the Kalman filter is studied as follows.

THEOREM 1. The state prediction and correction of algorithm (Eqs 9–12) equals the one of the extended Kalman filter if $\tilde{\mathbf{M}}_{k+1}$ is replaced by $\mathbf{K}_{k+1}^{\text{opt}}$ as stated in Eq. 24.

PROOF. ¹The state prediction (Eq. 9) obviously equals the one of the extended Kalman filter. Regarding the correct step (Eq. 11) it follows that

$$\hat{\mathbf{x}}_{k+1|k+1} = \mathbf{y}_{k+1} - \mathbf{K}_{k+1}^{\text{opt}} \mathbf{e}_{y_{k+1|k}} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{e}_{y_{k+1|k}} - \mathbf{K}_{k+1}^{\text{opt}} \mathbf{e}_{y_{k+1|k}}, \tag{25}$$

$$= \hat{\mathbf{x}}_{k+1|k} + (\mathbf{I}_{n \times n} - \mathbf{K}_{k+1}^{\text{opt}}) \mathbf{e}_{y_{k+1|k}} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}^{\text{Kal}} \mathbf{e}_{y_{k+1|k}}, \tag{26}$$

holds true if $\tilde{\mathbf{M}}_{k+1}$ is replaced by $\mathbf{K}_{k+1}^{\text{opt}}$ of (Eq. 24). As the introduced $\mathbf{K}_{k+1}^{\text{Kal}}$ of (Eq. 26) equals the Kalman filter gain

$$\mathbf{K}_{k+1}^{\text{Kal}} = \mathbf{I}_{n \times n} - \mathbf{K}_{k+1}^{\text{opt}} = \mathbf{I}_{n \times n} - \mathbf{R} (\mathbf{P}_{k+1|k} + \mathbf{R})^{-1}, \tag{27}$$

$$= (\mathbf{P}_{k+1|k} + \mathbf{R}) (\mathbf{P}_{k+1|k} + \mathbf{R})^{-1} - \mathbf{R} (\mathbf{P}_{k+1|k} + \mathbf{R})^{-1} \\ = \mathbf{P}_{k+1|k} (\mathbf{P}_{k+1|k} + \mathbf{R})^{-1}, \tag{28}$$

step (Eq. 26) and thus (Eq. 11) is identical to the correction step of the Kalman filter. ■

4. PARAMETER OPTIMIZATION OF THE KALMAN FILTER

The robustness of SVSF against model uncertainties is achieved by tuning the parameters of the smoothing boundary layer width [25]. However, also the Kalman filter gain can be made adaptive to achieve improved robustness [27]. In this section an adaptation law for the unknown a priori estimation error covariance $\mathbf{P}_{k+1|k}$ of the optimal gain (Eq. 24) is derived. The error covariance $\mathbf{P}_{k+1|k}$ should not be propagated based on the system model (like it is done usually in the field of Kalman filtering) as this would lead to prediction errors due to the imprecise model description. Instead $\mathbf{P}_{k+1|k}$ itself is required to be estimated.

¹The authors thank the anonymous reviewers of European Control Conference 2020 for the insightful comments and suggestions related to the proof of the theorem.

As according to Eq. 17 the state error covariance $\mathbf{P}_{k+1|k}$ is related to the output error covariance \mathbf{S}_{k+1} the expression

$$\hat{\mathbf{P}}_{k+1|k} = \hat{\mathbf{S}}_{k+1} - \mathbf{R}, \tag{29}$$

with

$$\hat{\mathbf{S}}_{k+1} = \frac{1}{N} \sum_{j=k-N+2}^{k+1} (\mathbf{e}_{y_{j|j-1}} \mathbf{e}_{y_{j|j-1}}^T), \tag{30}$$

$$= \frac{1}{N} \sum_{j=k-N+2}^{k+1} (\mathbf{e}_{x_{j|j-1}} \mathbf{e}_{x_{j|j-1}}^T + \mathbf{r}_j \mathbf{e}_{x_{j|j-1}}^T + \mathbf{e}_{x_{j|j-1}} \mathbf{r}_j^T + \mathbf{r}_j \mathbf{r}_j^T), \tag{31}$$

might be considered to gain information about $\mathbf{P}_{k+1|k}$. Estimating $\hat{\mathbf{S}}_{k+1}$ based on the innovation process $\mathbf{e}_{y_{k+1|k}}$ is common in the field of adaptive Kalman filtering, e.g., Refs. 27 and 28. From Eq. 31 it can be seen that if $\mathbf{e}_{x_{k+1|k}}$ is constant over N time steps, and \mathbf{r}_{k+1} is ergodic in the sense of $\frac{1}{N} \sum_{j=k-N+2}^{k+1} \mathbf{r}_j = 0$, and $\frac{1}{N} \sum_{j=k-N+2}^{k+1} \mathbf{r}_j \mathbf{r}_j^T = \mathbf{R}$ then $\hat{\mathbf{S}}_{k+1} - \mathbf{R}$ equals the a priori error covariance. Additional information about $\mathbf{P}_{k+1|k}$ can be obtained by considering the suboptimal gain $\mathbf{K}_{k+1}^{\text{sub}} = 0$. Replacing \mathbf{K}_{k+1} of Eq. 23 by the suboptimal gain $\mathbf{K}_{k+1}^{\text{sub}}$ gives the error covariance $\mathbf{P}_{k+1|k+1}^{\text{sub}} = \mathbf{R}$ of the suboptimal filter. As the filter with the optimal gain minimizes the MSE the upper bound

$$\text{tr}(\mathbf{P}_{k+1|k+1}) \leq \text{tr}(\mathbf{R}), \tag{32}$$

can be established. If it is assumed that $\mathbf{P}_{k+1|k+1} \approx \mathbf{P}_{k+1|k}$, which is the case for sufficient small sampling time then

$$\hat{\mathbf{P}}_{k+1|k} = \xi \mathbf{R}, \quad \xi \in [0, 1], \tag{33}$$

might be considered to gain information about $\mathbf{P}_{k+1|k}$. Due to the recursive nature of the filtering approach the relation of $\mathbf{P}_{k+1|k}$ to the previous value $\mathbf{P}_{k|k-1}$ might be considered. For small sampling time this can be roughly described by

$$\hat{\mathbf{P}}_{k+1|k} = \hat{\mathbf{P}}_{k|k-1}. \tag{34}$$

In order to find an estimation $\hat{\mathbf{P}}_{k+1|k}^{\star}$ that fits best to the established Eqs 29, 33, and 34 a weighted least squares (WLS) estimation problem is formulated. The importance of the individual equations is expressed by the scalar weights $\alpha, \beta, \gamma > 0$, which will be determined in a parameter optimization process. Using the vector operator “vec” the WLS problem

$$\mathbf{p}_{k+1}^{\star} = \arg \min_{\mathbf{P}_{k+1}} \|\mathbf{b}_{k+1} - \mathbf{A} \mathbf{p}_{k+1}\|_{\mathbf{W}}^2, \quad \hat{\mathbf{P}}_{k+1|k}^{\star} = \text{vec}^{-1}(\mathbf{p}_{k+1}^{\star}), \tag{35}$$

subject to

$$\mathbf{A} = \mathbf{I}_{n \times n} \otimes \begin{bmatrix} \mathbf{I}_{n \times n} \\ \mathbf{I}_{n \times n} \\ \mathbf{I}_{n \times n} \end{bmatrix}, \quad \mathbf{b}_{k+1} = \text{vec} \left(\begin{bmatrix} \hat{\mathbf{S}}_{k+1} - \mathbf{R} \\ \xi \mathbf{R} \\ \hat{\mathbf{P}}_{k|k-1}^{\star} \end{bmatrix} \right), \\ \mathbf{W} = \mathbf{I}_{n \times n} \otimes \begin{bmatrix} \alpha \mathbf{I}_{n \times n} & 0 & 0 \\ 0 & \beta \mathbf{I}_{n \times n} & 0 \\ 0 & 0 & \gamma \mathbf{I}_{n \times n} \end{bmatrix},$$

is considered. The solution of this WLS problem is

$$\hat{\mathbf{P}}_{k+1|k}^{\star} = \frac{1}{\alpha + \beta + \gamma} \left(\alpha(\hat{\mathbf{S}}_{k+1} - \mathbf{R}) + \beta\xi\mathbf{R} + \gamma\hat{\mathbf{P}}_{k|k-1}^{\star} \right), \quad (36)$$

which is a weighted sum of the solutions of the individual equations. The obtained solution (Eq. 36) is not guaranteed to be positive semidefinite. However, as $\mathbf{P}_{k+1|k}$ is a covariance matrix it must be at least positive semidefinite. In practice it will even be positive definite as otherwise the MSE would be zero. Incorporating the constraint $\mathbf{P}_{k+1} > 0$ into the WLS problem (Eq. 35) would lead to a nonlinear optimization problem. In order to keep the problem simple it is suggested to replace $\hat{\mathbf{P}}_{k+1|k}^{\star}$ by some $\hat{\mathbf{P}}_{k+1|k} > 0$ if $\hat{\mathbf{P}}_{k+1|k}^{\star}$ has non-positive eigenvalues. The modified error covariance $\hat{\mathbf{P}}_{k+1|k} > 0$ is

$$\hat{\mathbf{P}}_{k+1|k} = \mathbf{L}_{k+1}\mathbf{D}_{k+1}\mathbf{L}_{k+1}^T, \quad (37)$$

where \mathbf{L}_{k+1} diagonalizes $\hat{\mathbf{P}}_{k+1|k}^{\star}$ as

$$\mathbf{D}_{k+1}^{\star} = \mathbf{L}_{k+1}^T \hat{\mathbf{P}}_{k+1|k}^{\star} \mathbf{L}_{k+1}, \quad (38)$$

and matrix \mathbf{D}_{k+1} is a diagonal matrix with the i th diagonal element $(\mathbf{D}_{k+1})_{ii}$ defined by

$$(\mathbf{D}_{k+1})_{ii} = \begin{cases} (\mathbf{D}_{k+1}^{\star})_{ii} & \text{if } (\mathbf{D}_{k+1}^{\star})_{ii} \text{ is } > 0, \\ \eta \times (\mathbf{L}_{k+1}^T \mathbf{R} \mathbf{L}_{k+1})_{ii} & \text{if } (\mathbf{D}_{k+1}^{\star})_{ii} \text{ is } \leq 0, \end{cases}$$

where $i \in \{1, 2, \dots, n\}$. The scaling factor $\eta \in (0, 1]$ will be part of the parameter optimization process. The diagonalization (Eq. 38) can always be achieved as $\hat{\mathbf{P}}_{k+1|k}^{\star}$ is symmetric if a symmetric initial matrix $\hat{\mathbf{P}}_0$ is chosen. Finally, estimating the optimal Kalman filter gain $\mathbf{K}_{k+1}^{\text{opt}}$ based on $\hat{\mathbf{P}}_{k+1|k}$ leads to the adaptive Kalman filtering (A-KF) approach

$$\hat{\mathbf{x}}_{k+1|k} = \check{f}_k(\hat{\mathbf{x}}_{k|k}, \mathbf{u}_k), \quad (39)$$

$$\mathbf{e}_{y_{k+1|k}} = \mathbf{y}_{k+1} - \hat{\mathbf{x}}_{k+1|k}, \quad (40)$$

$$\hat{\mathbf{x}}_{k+1|k+1} = \mathbf{y}_{k+1} - \hat{\mathbf{K}}_{k+1}^{\text{opt}} \mathbf{e}_{y_{k+1|k}}, \quad (41)$$

$$\hat{\mathbf{K}}_{k+1}^{\text{opt}} = \mathbf{R}(\hat{\mathbf{P}}_{k+1|k} + \mathbf{R})^{-1}, \quad (42)$$

with

$$\hat{\mathbf{P}}_{k+1|k} = \mathbf{L}_{k+1}\mathbf{D}_{k+1}\mathbf{L}_{k+1}^T, \quad (43)$$

where \mathbf{L}_{k+1} diagonalizes $\hat{\mathbf{P}}_{k+1|k}^{\star}$ so that diagonalization \mathbf{D}_{k+1}^{\star} is achieved as

$$\mathbf{D}_{k+1}^{\star} = \mathbf{L}_{k+1}^T \hat{\mathbf{P}}_{k+1|k}^{\star} \mathbf{L}_{k+1}, \quad (44)$$

$$\hat{\mathbf{P}}_{k+1|k}^{\star} = \frac{1}{\alpha + \beta + \gamma} \left(\alpha(\hat{\mathbf{S}}_{k+1} - \mathbf{R}) + \beta\xi\mathbf{R} + \gamma\hat{\mathbf{P}}_{k|k-1} \right), \quad \alpha, \beta, \gamma > 0, \quad (45)$$

and the diagonal matrix \mathbf{D}_{k+1} is obtained from

$$(\mathbf{D}_{k+1})_{ii} = \begin{cases} (\mathbf{D}_{k+1}^{\star})_{ii} & \text{if } (\mathbf{D}_{k+1}^{\star})_{ii} \text{ is } > 0, \\ \eta \times (\mathbf{L}_{k+1}^T \mathbf{R} \mathbf{L}_{k+1})_{ii} & \text{if } (\mathbf{D}_{k+1}^{\star})_{ii} \text{ is } \leq 0, \end{cases} \quad (46)$$

with $i \in \{1, 2, \dots, n\}$ and factor $\eta \in (0, 1]$.

The adaptive filtering approach (Eqs 39–46) depends on the set

$$\mathcal{P} = \{N, \alpha, \beta, \gamma, \xi, \eta\}, \quad (47)$$

of tuning parameters. In order to optimize these parameters a training model is introduced. The filter is applied to the training model instead of the real system. The training model simulates the effects of model uncertainty occurring from the real unknown system by varying the system parameters of the known nominal model \check{f} . Let N_p be the number of system parameters and let $p_i^{(0)}$ be the nominal value of parameter i . Assume that a priori knowledge about the amount of model uncertainty is available meaning that parameter $p_i^{(0)}$ is required to be varied by q_i percent to account for the severity of model uncertainty. It is suggested to obtain the i th parameter of the training model $p_i^{(t)}$ by drawing a sample $p_i^{(t)} \sim U((1 - q_i/100)p_i^{(0)}, (1 + q_i/100)p_i^{(0)})$ from an uniform distribution. Repeating the procedure for all N_p system parameters forms one set of training parameters denoted as a training model. The optimization of the filter parameters based on the training model can be achieved as follows. Let $\bar{\mathbf{x}}_k$ be the states and $\bar{\mathbf{y}}_k$ be the measurements of the training model and let $\hat{\mathbf{x}}_{k|k}|\bar{\mathbf{y}}_k$ be the estimations of the filter using the nominal model but receiving the measurements from the training model $\bar{\mathbf{Y}}_k = \{\bar{\mathbf{y}}_0, \bar{\mathbf{y}}_1, \dots, \bar{\mathbf{y}}_k\}$ then the optimized parameters \mathcal{P}^{\star} are given as

$$\mathcal{P}^{\star} = \arg \min_{\mathcal{P}} J(\mathcal{P}) = \arg \min_{\mathcal{P}} \sum_k \|\bar{\mathbf{x}}_k - \hat{\mathbf{x}}_{k|k}|\bar{\mathbf{Y}}_k(\mathcal{P})\|^2, \quad (48)$$

where J denotes the costfunction. The optimization is a mixed integer optimization problem which can be solved by e.g., genetic approaches. In order to make the optimization process more reliable it is recommended to build several training models and optimize the filter parameters for each of them separately. The final optimized filter parameters can be obtained by taking the mean or median of the optimized parameters of the individual training models. The process of filter parameter optimization is illustrated in Figure 1.

5. STABILITY ANALYSIS OF THE ADAPTIVE KALMAN FILTER

As mentioned previously boundedness of the estimation error of SVSF approach (Eqs 3–7) using smoothing boundary layer has never been proven. However, if the adaptive Kalman filter is used instead ($\hat{\mathbf{M}}_{k+1}$ is replaced by $\hat{\mathbf{K}}_{k+1}^{\text{opt}}$) the boundedness of the estimation error can be proven as follows.

THEOREM 2. The estimation error of the a posteriori estimation

$$\hat{\mathbf{x}}_{k+1|k+1} = \mathbf{y}_{k+1} - \hat{\mathbf{K}}_{k+1}^{\text{opt}} \mathbf{e}_{y_{k+1|k}}, \quad (49)$$

of the adaptive Kalman filtering approach (Eqs 39–46) is bounded by

$$\|\mathbf{e}_{\mathbf{x}_{k+1|k+1}}\| \leq \sqrt{\frac{a_2^2}{4\epsilon_1\epsilon_2}} \mathbf{R} + a_1, \quad (50)$$

if the measurement noise is bounded by $\|\mathbf{r}_k\| < a_1$ and the following conditions are fulfilled

$$\mathbf{\Omega}_k = \hat{\mathbf{P}}_{k+1|k} - \epsilon_1 \mathbf{e}_{y_{k+1|k}} \mathbf{e}_{y_{k+1|k}}^T \geq 0, \quad (51)$$

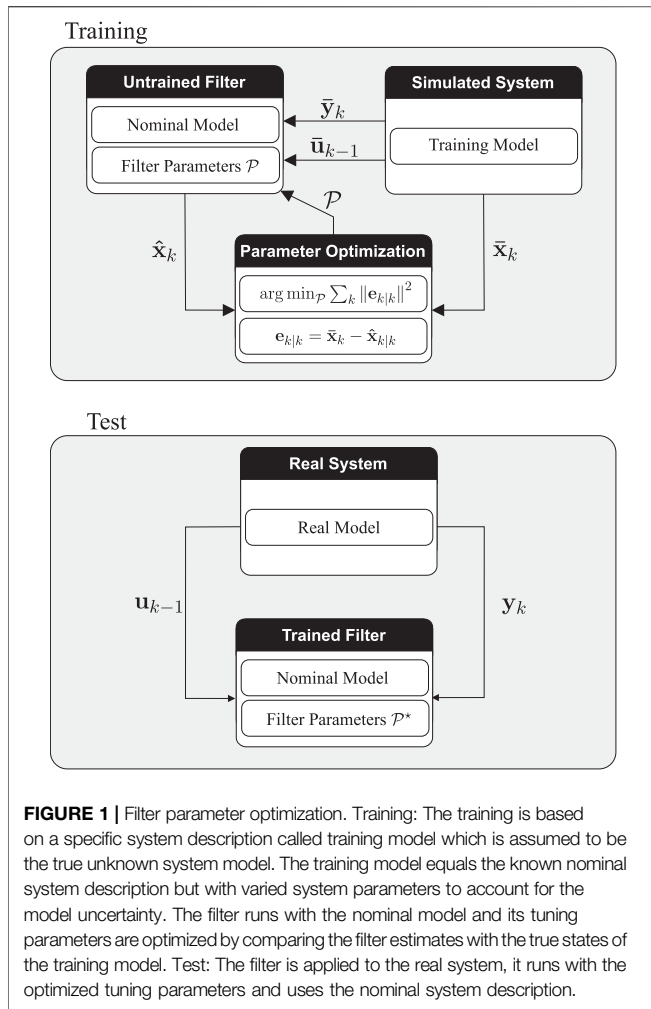


FIGURE 1 | Filter parameter optimization. Training: The training is based on a specific system description called training model which is assumed to be the true unknown system model. The training model equals the known nominal system description but with varied system parameters to account for the model uncertainty. The filter runs with the nominal model and its tuning parameters are optimized by comparing the filter estimates with the true states of the training model. Test: The filter is applied to the real system, it runs with the optimized tuning parameters and uses the nominal system description.

$$\epsilon_1 \mathbf{e}_{y_{k+1|k}}^T (\mathbf{R} + \mathbf{\Omega}_k)^{-1} \mathbf{e}_{y_{k+1|k}} \neq 1, \tag{52}$$

$$\lambda_{\min}\{(\mathbf{R} + \mathbf{\Omega}_k)^{-1}\} > \epsilon_2, \tag{53}$$

$$a_2 > \|\mathbf{R} + \mathbf{\Omega}_k\|, \tag{54}$$

where ϵ_1, ϵ_2 are sufficient small positive constants, a_1, a_2 are sufficient large positive constants.

PROOF. From Eq. 51 it can be concluded that $\hat{\mathbf{P}}_{k+1|k}$ can be written as

$$\hat{\mathbf{P}}_{k+1|k} = \epsilon_1 \mathbf{e}_{y_{k+1|k}} \mathbf{e}_{y_{k+1|k}}^T + \mathbf{\Omega}_k, \tag{55}$$

with $\mathbf{\Omega}_k \geq 0$. As known from Eq. 13 the a posteriori estimation error is

$$\mathbf{e}_{x_{k+1|k+1}} = \hat{\mathbf{K}}_{k+1}^{\text{opt}} \mathbf{e}_{y_{k+1|k}} - \mathbf{r}_{k+1}. \tag{56}$$

With the definition $\tilde{\mathbf{R}}_k = \mathbf{R} + \mathbf{\Omega}_k > 0$ the estimated optimal filter gain can be written as

$$\hat{\mathbf{K}}_{k+1}^{\text{opt}} \mathbf{e}_{y_{k+1|k}} = \mathbf{R} \left(\tilde{\mathbf{R}}_k + \epsilon_1 \mathbf{e}_{y_{k+1|k}} \mathbf{e}_{y_{k+1|k}}^T \right)^{-1} \mathbf{e}_{y_{k+1|k}}. \tag{57}$$

Applying the Sherman-Morrison-Woodbury formulur (Fact 2.16.3 [29]) in combination with assumption (Eq. 52) gives

$$\begin{aligned} \left(\tilde{\mathbf{R}}_k + \epsilon_1 \mathbf{e}_{y_{k+1|k}} \mathbf{e}_{y_{k+1|k}}^T \right)^{-1} \mathbf{e}_{y_{k+1|k}} &= \tilde{\mathbf{R}}_k^{-1} \mathbf{e}_{y_{k+1|k}} - \frac{\tilde{\mathbf{R}}_k^{-1} \epsilon_1 \mathbf{e}_{y_{k+1|k}} \mathbf{e}_{y_{k+1|k}}^T \tilde{\mathbf{R}}_k^{-1} \mathbf{e}_{y_{k+1|k}}}{1 + \mathbf{e}_{y_{k+1|k}}^T \tilde{\mathbf{R}}_k^{-1} \epsilon_1 \mathbf{e}_{y_{k+1|k}}}, \\ &= \frac{\tilde{\mathbf{R}}_k^{-1} \mathbf{e}_{y_{k+1|k}}}{1 + \epsilon_1 \mathbf{e}_{y_{k+1|k}}^T \tilde{\mathbf{R}}_k^{-1} \mathbf{e}_{y_{k+1|k}}}. \end{aligned} \tag{58}$$

Using the submultiplicative Frobenius norm (Proposition 9.3.5 [29]) and Eq. 58, Eq. 57 can be written as

$$\|\hat{\mathbf{K}}_{k+1}^{\text{opt}} \mathbf{e}_{y_{k+1|k}}\| \leq \frac{\|\mathbf{R}\| \|\tilde{\mathbf{R}}_k^{-1}\| \|\mathbf{e}_{y_{k+1|k}}\|}{1 + \epsilon_1 \mathbf{e}_{y_{k+1|k}}^T \tilde{\mathbf{R}}_k^{-1} \mathbf{e}_{y_{k+1|k}}}. \tag{59}$$

Based on the minimum eigenvalue of $\tilde{\mathbf{R}}_k^{-1}$ the inequality

$$\epsilon_1 \lambda_{\min}\{\tilde{\mathbf{R}}_k^{-1}\} \|\mathbf{e}_{y_{k+1|k}}\|^2 \leq \epsilon_1 \mathbf{e}_{y_{k+1|k}}^T \tilde{\mathbf{R}}_k^{-1} \mathbf{e}_{y_{k+1|k}}, \tag{60}$$

can be established (Lemma 8.4.1 [29]). Using Eq. 60 and assumption Eq. 53 an upper bound of Eq. 59 is obtained as

$$\|\hat{\mathbf{K}}_{k+1}^{\text{opt}} \mathbf{e}_{y_{k+1|k}}\| \leq \frac{\|\mathbf{R}\| \|\tilde{\mathbf{R}}_k^{-1}\| \|\mathbf{e}_{y_{k+1|k}}\|}{1 + \epsilon_1 \epsilon_2 \|\mathbf{e}_{y_{k+1|k}}\|^2}. \tag{61}$$

Unfortunately, $\tilde{\mathbf{R}}_k^{-1}$ implicitly depends on $\mathbf{e}_{y_{k+1|k}}$ so that assumption Eq. 54 has to be taken into consideration leading to

$$\|\hat{\mathbf{K}}_{k+1}^{\text{opt}} \mathbf{e}_{y_{k+1|k}}\| \leq \frac{a_2 \|\mathbf{R}\| \|\mathbf{e}_{y_{k+1|k}}\|}{1 + \epsilon_1 \epsilon_2 \|\mathbf{e}_{y_{k+1|k}}\|^2}. \tag{62}$$

Based on the derivative of Eq. 62 with respect to $\|\mathbf{e}_{y_{k+1|k}}\|$ it can be shown that the maximum of the upper bound is achieved for $\|\mathbf{e}_{y_{k+1|k}}\| = \sqrt{\frac{1}{\epsilon_1 \epsilon_2}}$ which leads to

$$\|\hat{\mathbf{K}}_{k+1}^{\text{opt}} \mathbf{e}_{y_{k+1|k}}\| \leq \sqrt{\frac{a_2^2}{4\epsilon_1 \epsilon_2}} \|\mathbf{R}\|. \tag{63}$$

Then (Eq. 50) is proven by applying the triangle inequality on Eq. 56.

6. NUMERICAL EXAMPLE

In this section state estimation and control of a chemical plant is considered in order to evaluate the performance of original SVSF and the adaptive Kalman filtering variant. According to Ref. 30, Ref. 31 a species *A* reacts in a continuous stirred tank reactor. The dynamics of the effluent flow concentration $C_A = x_1$ of species *A* and the reactor temperature $T = x_2$ can be described by the time-continuous model

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x} \end{bmatrix} = \begin{bmatrix} \frac{q}{V}(C_{Af} - x_1) - k_0 x_1 \exp\left(-\frac{E}{R x_2}\right) \\ a + \frac{UA}{V \rho C_p} (u + T_c^{eq} - x_2) \\ a = \frac{q}{V}(T_f - x_2) + \frac{-\Delta H k_0 x_1}{\rho C_p} \exp\left(-\frac{E}{R x_2}\right) \end{bmatrix}, \quad (64)$$

with measurements y , and control variable z defined as

$$y = x + r, \quad z = x_1. \quad (65)$$

The parameters of the system are shown in **Table 2**. The input of the system is the change of the coolant stream temperature $u = \Delta T_c$ related to the nominal value T_c^{eq} . An input saturation of $|u| \leq 50$ K is considered. The system is simulated based on Euler method using a sampling time of 0.1 s. The resulting discrete-time measurement model is

$$y_k = x_k + r_k, \quad (66)$$

where the zero-mean, white noise r_k is Gaussian with covariance

$$R = \begin{bmatrix} \sigma_{C_A}^2 & 0 \\ 0 & \sigma_T^2 \end{bmatrix}, \quad \sigma_{C_A}^2 = 0.8 \frac{\text{mol}^2}{\text{m}^6}, \quad \sigma_T^2 = 0.5 \text{K}^2.$$

For the initial state values $x_1(t_0) = 0.875$ mol/l, and $x_2(t_0) = 325$ K are considered.

6.1. Filter Training Process

In the following the optimization of the tuning parameters of A-KF and SVSF approach is considered. As explained in **Section 4** variation of the nominal system parameters is required to build the training model. To account for the model uncertainty a variation of 20% of the nominal parameters is considered which is assumed as a priori known. Consequently, the i th parameter of the training model $p_i^{(t)}$ is obtained from the uniform distribution $p_i^{(t)} \sim U((1 + 0.2)p_i^{(0)}, (1 - 0.2)p_i^{(0)})$. Repeating the step for all parameters $i \in \{1, 2, \dots, 10\}$ forms one training model. The parameter T_c^{eq} is not varied as it is required to be precisely known for control. In **Table 3** an overview of the true and nominal system parameters and the parameters used to build the training model is given. In order to account for the different combination and variation of the system parameters three training models are build denoted as Training I, Training II, and Training III.

For each set the tuning parameters of the filters are optimized separately. The parameters of A-KF approach required to be optimized are given by **Eq. 47**. For the SVSF approach the boundary layer widths Ψ_{11} , Ψ_{22} and convergence rates Φ_{11} , Φ_{22} of the first and second state are optimized. The optimization is achieved using “genetic algorithm” of MATLAB with default settings. During optimization only one realization of measurement noise is considered so that the costfunction does not vary due to the noise.

The results of optimization are shown in **Table 4**. The A-KF approach requires at least 60 times more computational time

TABLE 2 | System parameter description [30].

Description	Parameter	Unit
Tank volume	V	l
Feed flow rate	q	l/min
Feed concentration	C_{Af}	mol/l
Feed temperature	T_f	K
Density	ρ	g/l
Enthalpy	$-\Delta H$	J/mol
Exponential factor	$\frac{E}{R}$	K
Frequency factor	k_0	min ⁻¹
Heat transfer characteristic	UA	J/minK
Specific heat	C_p	J/gK
Coolant flow temperature	T_c^{eq}	K

TABLE 3 | System parameters of nominal, real, and training model.

Parameter	Nominal	Real	Training
V	100	92	[80, 120]
Q	100	94	[80, 120]
C_{Af}	1	0.98	[0.8, 1.2]
T_f	350	354	[280, 420]
ρ	1,000	953	[800, 1,200]
$-\Delta H$	$5e^4$	$4.91e^4$	$[4e^4, 6e^4]$
$\frac{E}{R}$	8,750	8,893	[7,000, 10,500]
k_0	$7.2e^{10}$	$7.68e^{10}$	$[5.76e^{10}, 8.64e^{10}]$
UA	$5e^4$	$4.68e^4$	$[4e^4, 6e^4]$
C_p	0.239	0.299	[0.191, 0.287]
T_c^{eq}	300	300	300

Interval $[a,b]$ denotes: Parameter for training equals a sample u drawn from uniformly distribution $U(a,b)$, i.e., $u \sim U(a,b)$.

TABLE 4 | Computational effort of training process.

	Training I		Training II		Training III	
	A-KF	SVSF	A-KF	SVSF	A-KF	SVSF
GA iterations	4,795	4,330	6,049	3,813	6,106	4,095
GA convergence	Achieved	Achieved	Achieved	Achieved	Achieved	Achieved
Min. value of costfunction	0.26877	0.49976	0.41247	0.74274	0.37931	0.56349
Total number of FLOPS	$74867e^6$	$779e^6$	$94446e^6$	$686e^6$	$95337e^6$	$737e^6$
Training time [min]	173.24	2.48	215.23	2.37	208.60	2.65

GA: Genetic Algorithm of MATLAB, GA iterations: Calls of costfunction, GA convergence: Change of costfunction falls below tolerance, Min. value of costfunction equals $J(P^*)$ of **Eq. 48**, FLOPS: floating point operations, Hardware: 4xCPU@3.7GHz with 8 GB memory.

for the optimization than SVSF. The SVSF algorithm is computationally more efficient and in addition it requires less parameters to be optimized. However, the achieved minimal value of the costfunction is always lower in case of A-KF.

6.2. Open Loop Case

In the open loop case the step functions $u = 5$ K and $u = -5$ K are applied to the system and the resulting step responses are

considered for state estimation. The system behavior is simulated over a time horizon of 6 min.

The performance values of A-KF and SVSF approach are shown in **Table 5**. The filters are applied based on four sets of optimized parameters. Three sets result from Training I, Training II, and Training III. The fourth set denoted as Training I-III considers the mean values of the optimized parameters of the other three sets. As all states are measured the MSE of the measurements is considered also. In order to account for the effect of different noise realizations the results are obtained by simulating the step response 100 times and taking the mean value and variance of the squared estimation error. The filters are applied to the same measurements with same noise realizations. From the results it can be seen that both A-KF and SVSF estimations are more precise than the measurements. In comparison to SVSF the A-KF approach achieves better estimation performance for all considered sets of optimized parameters and both step responses. For one specific run the step responses and the corresponding state estimations are visualized in **Figures 2** and **3**. The computational time required to generate **Table 5** (to simulate both step responses 100 times and apply 8 filters each time) is 18.90 min on a 4xCPU@3.7Ghz with 8 GB memory. The computational time required to apply only one filter on one specific step response is 0.11 s in case of SVSF and 1.49 s in case of A-KF. Both values are far less than the 6 min of simulated system behavior.

6.3. Closed Loop Case

In the closed loop case the effluent flow concentration C_A of the chemical plant is controlled. The controller relies on the filter estimates. Consequently, the estimation performance as well as the control performance dependent on the filters is studied.

In order to achieve reference tracking of the control variable C_A the super-twisting sliding mode approach [32] is applied. By introducing the estimated tracking error

$$\hat{e}_r = z_r - \hat{z}, \tag{67}$$

with reference value z_r , the sliding variable can be calculated as

$$\hat{s} = \hat{e}_r + \lambda \hat{e}_r, \quad \lambda > 0. \tag{68}$$

Then reference tracking can be achieved by applying the super-twisting approach [32]

$$u = -k_1 |\hat{s}|^{\frac{1}{2}} \text{sgn}(\hat{s}) - k_2 \int \text{sgn}(\hat{s}) dt \tag{69}$$

The controller parameters are chosen as $k_1 = 0.001$ and $k_2 = 25$ based on trial and error. For the sliding dynamics $\lambda = 0.08$ is considered. The simulated time horizon is 20 min with reference values defined by

TABLE 5 | Estimation performance in open loop case (100 realizations).

		Training I		Training II		Training III		Training I-III		Meas
		A-KF	SVSF	A-KF	SVSF	A-KF	SVSF	A-KF	SVSF	—
$u = +5 \text{ K}$	μ_{e_x}	0.194	0.297	0.198	0.509	0.200	0.295	0.197	0.329	0.889
	$\sigma_{e_x}^2$	$7.63e^{-5}$	$7.10e^{-5}$	$7.43e^{-5}$	$2.44e^{-4}$	$8.82e^{-5}$	$1.00e^{-4}$	$8.27e^{-5}$	$1.21e^{-4}$	$2.66e^{-4}$
$u = -5 \text{ K}$	μ_{e_x}	0.114	0.276	0.117	0.503	0.117	0.279	0.117	0.314	0.892
	$\sigma_{e_x}^2$	$7.65e^{-5}$	$8.50e^{-5}$	$7.12e^{-5}$	$2.59e^{-4}$	$8.86e^{-5}$	$1.20e^{-4}$	$8.45e^{-5}$	$1.37e^{-4}$	$2.67e^{-4}$

Meas., Measurements, μ_{e_x} : Mean of $e_x^T e_x$ (squared estimation error), $\sigma_{e_x}^2$: Variance of $e_x^T e_x$ (squared estimation error).

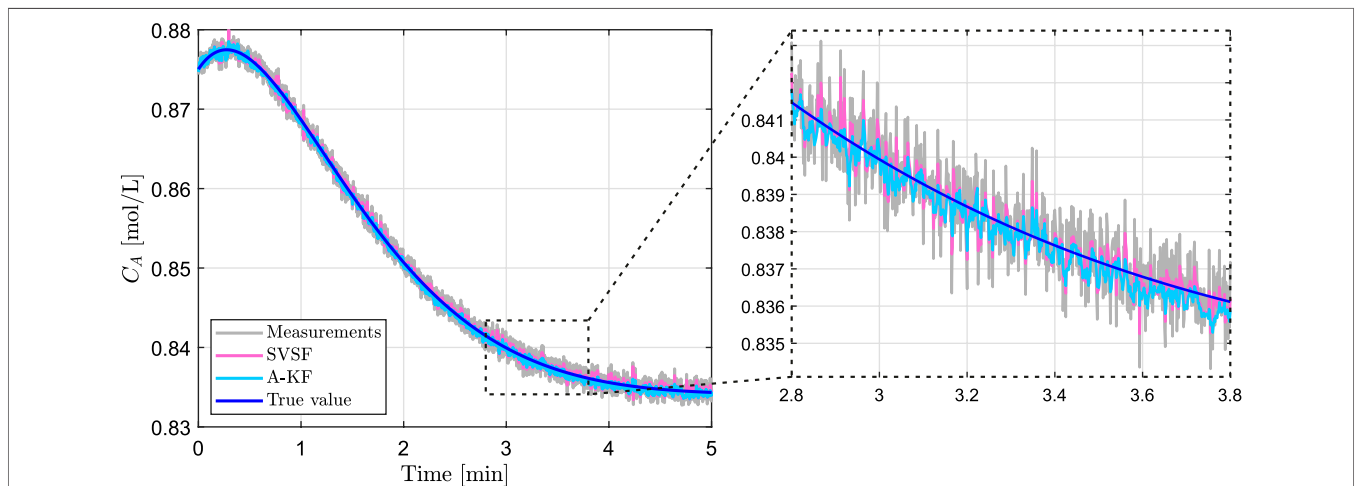


FIGURE 2 | Open loop response of step input $u = +5 \text{ K}$ (Estimation errors of shown realization: $\mu_{e_x}^{A-KF} = 0.193$, $\mu_{e_x}^{SVSF} = 0.296$, $\mu_{e_x}^{Meas.} = 0.892$).

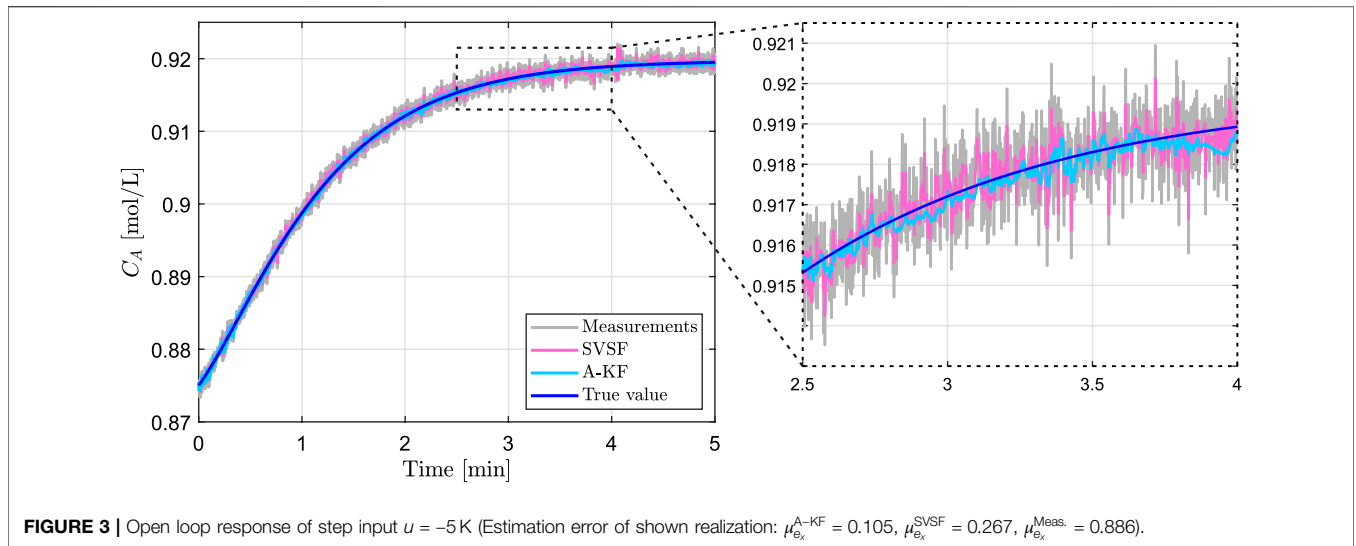


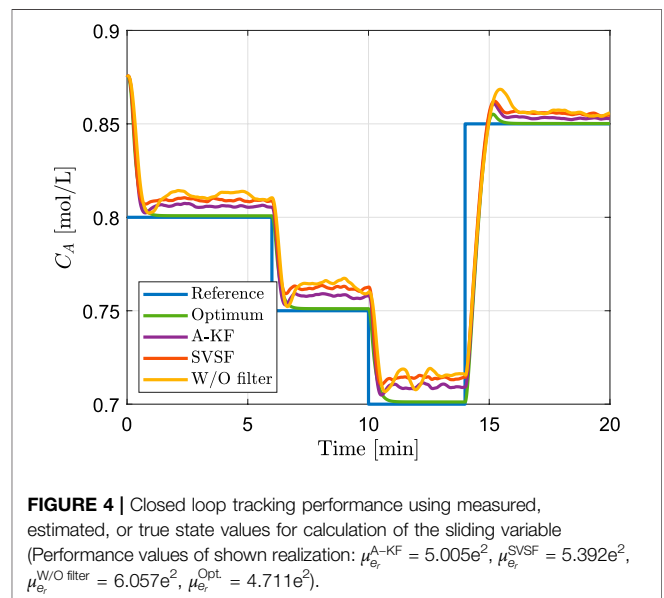
TABLE 6 | Control and estimation performance in closed loop case (100 realizations).

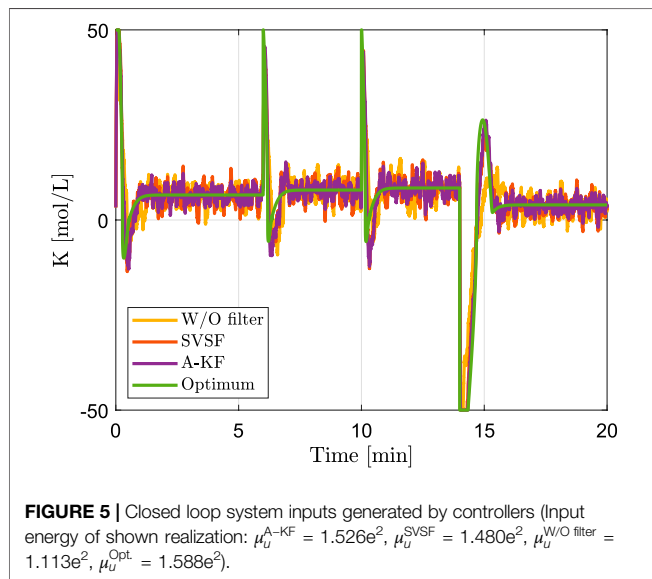
	Training I		Training II		Training III		Training I-III		W/O	Opt
	A-KF	SVSF	A-KF	SVSF	A-KF	SVSF	A-KF	SVSF	Filter	—
μ_{e_t}	4.999e ²	5.420e ²	5.006e ²	5.884e ²	5.006e ²	5.518e ²	5.005e ²	5.562e ²	6.229e ²	4.711e ²
$\sigma_{e_t}^2$	31.512	46.292	31.441	62.391	31.625	58.460	31.787	60.623	1.939e2	0
μ_{e_x}	0.221	0.327	0.222	0.520	0.225	0.322	0.222	0.352	0.890	0
$\sigma_{e_x}^2$	2.38e ⁻⁵	3.55e ⁻⁵	2.44e ⁻⁵	8.70e ⁻⁵	2.43e ⁻⁵	4.36e ⁻⁵	2.37e ⁻⁵	5.12e ⁻⁵	9.07e ⁻⁵	0
μ_u	1.557e ²	1.495e ²	1.551e ²	1.369e ²	1.559e ²	1.460e ²	1.554e ²	1.458e ²	1.086e ²	1.588e ²
σ_u^2	14.383	4.7964	15.022	7.301	15.397	5.928	14.968	6.100	8.758	0

Opt., Optimum; μ_{e_t} : Mean of e_t^2 (squared tracking error), $\sigma_{e_t}^2$: Variance of e_t^2 (squared tracking error), μ_{e_x} : Mean of $e_x^T e_x$ (squared estimation error), $\sigma_{e_x}^2$: Variance of $e_x^T e_x$ (squared estimation error), μ_u : Mean of u^2 (input energy), σ_u^2 : Variance of u^2 (input energy).

$$z_r = \begin{cases} 0.8 \text{ mol/l} & \text{if } t \leq 6 \text{ min,} \\ 0.75 \text{ mol/l} & \text{if } 6 \text{ min} < t \leq 10 \text{ min,} \\ 0.7 \text{ mol/l} & \text{if } 10 \text{ min} < t \leq 14 \text{ min,} \\ 0.85 \text{ mol/l} & \text{if } 14 \text{ min} < t \leq 20 \text{ min,} \end{cases} \quad (70)$$

where $t \in [0 \text{ min}, 20 \text{ min}]$. Related to the control performance the following four cases are studied. The sliding variable (Eq. 68) is calculated using the estimations of A-KF approach or the estimations of SVSF approach or the measurements denoted as “W/O filters” or the true states denoted as “Optimum”. For the filters the optimized filter parameters of the different training models are considered. The control of the plant over the time horizon of 20 min is simulated 100 times to account for the noise realizations. The measurements used by the filters have same noise realizations. The performance values are shown in Table 6. The best control performance can be achieved based on the true states which are noise-free. Control based on estimations of A-KF or SVSF achieves better performance than applying the controller directly on the noisy measurements. In comparison to SVSF approach A-KF method achieves better reference tracking and more precise state estimation





for all considered training models. The control variable and the input values of one specific realization are visualized in **Figures 4 and 5**. The computational time required to generate **Table 6** (to simulate 100 runs of 10 closed loop systems in parallel) is 34.54 min on a 4xCPU@3.7Ghz with 8 GB memory. The computational time required to simulate the closed loop behavior only once based on one controller is 0.45 s if the controller is fed by the estimations of SVSF and 5.24 s if the controller is fed by the estimations of A-KF. This is in both cases far less than the time horizon of 20 min over which the system behavior is simulated.

REFERENCES

- Kitanidis PK. Unbiased minimum-variance linear state estimation. *Automatica* (1987) 23:775–8. doi:10.1016/0005-1098(87)90037-9.
- Gilljins S, De Moor B. Unbiased minimum-variance input and state estimation for linear discrete-time systems. *Automatica* (2007) 43:111–6. doi:10.1016/j.automatica.2006.08.002.
- Anderson BD, Moore JB. *Optimal filtering*. Englewood Cliffs, NJ: Prentice-Hall (1979) 256 p.
- Hmida FB, Khémiri K, Ragot J, Gossa M. Three-stage Kalman filter for state and fault estimation of linear stochastic systems with unknown inputs. *J Franklin Inst* (2012) 349:2369–88. doi:10.1016/j.jfranklin.2012.05.004.
- Dong Z, You Z. Finite-horizon robust Kalman filtering for uncertain discrete time-varying systems with uncertain-covariance white noises. *IEEE Signal Process Lett* (2006) 13:493–6. doi:10.1109/LSP.2006.873148.
- Hassibi B, Sayed AH, Kailath T. *Indefinite quadratic estimation and control: a unified approach to H_2 and H_∞ theories* Society for Industrial and Applied Mathematics. Philadelphia, PA (1999) p. 572.
- Blom HAP, Bar-Shalom Y. The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Trans Automat Contr*. (1988) 33:780–3. doi:10.1109/9.1299.
- Martino L, Read J, Elvira V, Louzada F. Cooperative parallel particle filters for online model selection and applications to urban mobility. *Digit Signal Process* (2017) 60:172–85. doi:10.1016/j.dsp.2016.09.011.
- Habibi S. The smooth variable structure filter. *Proc IEEE* (2007). 95:1026–59. doi:10.1109/JPROC.2007.893255.
- Gadsden SA, Song Y, Habibi SR. Novel model-based estimators for the purposes of fault detection and diagnosis. *IEEE ASME Trans Mechatron* (2013) 18:1237–49. doi:10.1109/TMECH.2013.2253616.

7. CONCLUSION

In this paper the relation of smooth variable structure filter (SVSF) and Kalman filter has been studied. An adaptive Kalman filter (A-KF) was derived from the SVSF approach. Sufficient conditions for the boundedness of the estimation error of A-KF approach have been formulated. The simulation results show that beside the SVSF approach also an adaptive Kalman filter can be used to achieve robust estimation performance.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

The development and theoretical description of the adaptive Kalman filter approach was undertaken by MS and DS. The numerical simulation was designed and realized by MS. The manuscript was drafted by MS and reviewed by DS.

ACKNOWLEDGMENTS

The authors want to thank the reviewers for their valuable comments and suggestions.

- Kim T, Wang Y, Fang H, Sahinoglu Z, Wada T, Hara S, et al. Model-based condition monitoring for lithium-ion batteries. *J Power Sources* (2015) 295: 16–27. doi:10.1016/j.jpowsour.2015.03.184.
- Qiao HH, Attari M, Ahmed R, Delbari A, Habibi S, Shoa T. Reliable state of charge and state of health estimation using the smooth variable structure filter. *Contr Eng Pract* (2018) 77:1–14. doi:10.1016/j.conengprac.2018.04.015.
- Gadsden SA, Habibi SR, Kirubarajan T. A novel interacting multiple model method for nonlinear target tracking. In: FUSION 2010 : 13th international conference on information fusion; 2010 Jul 26 - 29; Edinburgh, Scotland. Piscataway, NJ: IEEE.
- Attari M, Gadsden SA, Habibi SR. Target tracking formulation of the SVSF as a probabilistic data association algorithm. In: The 2013 American control conference; 2013 June 17 - 19; Washington, DC; IEEE, Piscataway, NJ. p. 6328–32. doi:10.1109/ACC.2013.6580830.
- Attari M, Luo Z, Habibi S. An SVSF-based generalized robust strategy for target tracking in clutter. *IEEE Trans Intell Transport Syst* (2016) 17:1381–92. doi:10.1109/TITS.2015.2504331.
- Luo Z, Attari M, Habibi S, Mohrenschildt MV. Online multiple maneuvering vehicle tracking system based on multi-model smooth variable structure filter. *IEEE Trans Intell Transport Syst* (2020) 21:603–16. doi:10.1109/TITS.2019.2899051.
- Demim F, Nemra A, Louadj K. Robust SVSF-SLAM for unmanned vehicle in unknown environment. *IFAC-PapersOnLine* (2016) 49:386–94. doi:10.1016/j.ifacol.2016.10.585.
- Allam A, Tadjine M, Nemra A, Kobzili E. Stereo vision as a sensor for SLAM based smooth variable structure filter with an adaptive boundary layer width. In: ICSC 2017, 6th International conference on systems and control. 2017 May 7–9; Batna, Algeria; Piscataway, NJ: IEEE. p. 14–20. doi:10.1109/ICoSC.2017.7958700.

19. Liu Y, Wang C. A FastSLAM based on the smooth variable structure filter for UAVs. In: 15th international conference on ubiquitous robots (UR); June 26-30, 2018; Hawaii Convention Center; Piscataway, NJ IEEE (2018) p. 591–6. doi:10.1109/URAI.2018.8441876.
20. Ahmed RM, Sayed MAE, Gadsden SA, Habibi SR. Fault detection of an engine using a neural network trained by the smooth variable structure filter. In: IEEE international conference on control applications (CCA); Denver, CO; 2011 Sep 28-30; Piscataway, NJ IEEE (2011) p. 1190–6. doi:10.1109/CCA.2011.6044515.
21. Al-Shabi M, Habibi S. Iterative smooth variable structure filter for parameter estimation. *ISRN Signal Processing* (2011) 2011:1. doi:10.5402/2011/725108.
22. Spiller M, Bakhshande F, Söffker D. The uncertainty learning filter: a revised smooth variable structure filter. *Signal Process* (2018) 152:217–26. doi:10.1016/j.sigpro.2018.05.025.
23. Longbin M, Xiaoquan S, Yiyu Z, Kang SZ, Bar-Shalom Y. Unbiased converted measurements for tracking. *IEEE Trans Aero Electron Syst* (1998) 34:1023–7. doi:10.1109/7.705921.
24. Gadsden SA, Habibi SR. A new form of the smooth variable structure filter with a covariance derivation. In: CDC 2010 : 49th IEEE conference on decision and control 2010 Dec 15 - 17; Atlanta, GA; Piscataway, NJ: IEEE (2010) p. 7389–94. doi:10.1109/CDC.2010.5717397.
25. Gadsden SA, El Sayed M, Habibi SR. Derivation of an optimal boundary layer width for the smooth variable structure filter. In: 2011 American Control Conference - ACC 2011; San Francisco, CA; 2011 Jun 29 - Jul 1; Piscataway, NJ: IEEE (2011) p. 4922–7. doi:10.1109/ACC.2011.5990970.
26. Tian Y, Suwoyo H, Wang W, Li L. An ASVSF-SLAM algorithm with time-varying noise statistics based on MAP creation and weighted exponent. *Math Probl Eng*. (2019) 2019:1. doi:10.1155/2019/2765731.
27. Hide C, Moore T, Smith M. Adaptive Kalman filtering for low-cost INS/GPS. *J Navig*. (2003) 56:143–52. doi:10.1017/S0373463302002151.
28. Yang Y, Gao W. An optimal adaptive Kalman filter. *J Geodes*. (2006) 80: 177–83. doi:10.1007/s00190-006-0041-0.
29. Bernstein DS. *Matrix mathematics: theory, facts, and formulas* Princeton, NJ: Princeton University Press (2009) 1184 p.
30. Magni L, Nicolao GD, Magnani L, Scattolini R. A stabilizing model-based predictive control algorithm for nonlinear systems. *Automatica* (2001) 37: 1351–62. doi:10.1016/S0005-1098(01)00083-8.
31. Seborg DE, Mellichamp DA, Edgar TF, Doyle FJ, III. *Process dynamics and control* New York, NY: John Wiley & Sons (2010) p. 1085–6.
32. Shtessel Y, Edwards C, Fridman L, Levant A. *Sliding mode control and observation*. New York: Springer (2014) 356 p.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Spiller and Söffker. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.