2018 26th Mediterranean Conference on Control and Automation (MED)
Zadar, Croatia, June 19-22, 2018

Desensitized Filtering for Systems with Uncertain Parameters and Noise Correlation*

Jaroslav Tabaček[1] and Vladimír Havlena[1]

Abstract— **This paper introduces estimation algorithms for systems with uncertain parameters and correlated noises. The algorithms are derived using the standard Kalman filter for correlated noises and the desensitized filtering technique for systems with uncertain parameters. A general algorithm and its special case are proposed. The latter updates statistics with explicit expressions, which makes it simpler and faster. The extended forms of the algorithms, which can be used for nonlinear systems, are also introduced. The developed algorithm is tested on an example, where the importance of the noise correlation information is shown.**

I. INTRODUCTION

System descriptions often include uncertain parameters as a consequence of creating models that significantly simplify the structure of the models of real systems, but their behavior still sufficiently approximates the behavior of real systems. The uncertain parameters in simplified models are the parameters that are assumed to be constant, but their true value cannot be described exactly. However, it can usually be described by a probability distribution. The models with uncertain parameters can be used in conventional state estimation algorithms, but the uncertainty in parameters causes that optimality and stability of these algorithms cannot be guaranteed. Therefore, special algorithms are used for the state estimation of these systems.

Information about the correlation of measurement and process noise is often neglected in the development of models and estimation algorithms. However, the cost of adding such information is low, and it can significantly improve the estimation. For example, the noise correlation needs to be considered in the models which are created by discretization of a continuous-time system with discrete-time measurements. When a control algorithm is developed for such a system, it should use the latest information about the system state to increase the accuracy of the control. Therefore, the asynchronous sampling of the control input and the measurements are used. Such sampling leads to the discretization of the process and measurement equation, which creates the noise correlation in the system.

There are several state estimation methods available for systems with uncertain parameters. The first option is to use the Extended Kalman Filter (EKF) with the state vector augmented by uncertain parameters [1]. This way, the uncertainty is taken into account, and moreover, the parameter vector is estimated. The downside of this method is that the EKF can estimate infeasible values of the parameters, which will cause linearization of the system in an infeasible point and consequently a crash of the algorithm.

The next option is to use the Schmidt-Kalman Filter (SKF) [2], [1] or its variation the 'Consider' Kalman Filter (CKF) [3]. Similarly to the EKF, this method augments the state vector by the parameters. However, in this case, the Kalman filter is used in the reduced-state form where the parameter statistics is not updated, but it is used for the update of the state statistics. Several similar reduced-state Kalman filters were developed, and their summary can be found in [4]. The CKF and SKF are derived with the assumption of uncorrelated noises, and introducing the noise correlation to the derivation of the CKF increases algorithm complexity significantly.

The class of observers called the unknown input observers [5], [6] can also be used for the state estimation of the systems with uncertainty, but these observers are based on the assumption of a deterministic system. An effective method for the state estimation of the stochastic systems with uncertainty is the robust H∞ filter [7], [8] or other effective robust Kalman filter methods [9]. These methods are good at compensating the system uncertainty but they tent to have low sensitivity to state deviations.

Another robust algorithm is the Desensitized Kalman Filter (DKF) [10]. This method includes uncertainty information into the Kalman filter algorithm by adding the information about state estimation error sensitivity to the particular parameters. The state is estimated in a way which minimizes this sensitivity and the trace of an estimation error covariance at the same time. In this way, even a large error in the parameter value causes only a small state estimation error. The special case of DKF was proposed in [11], in which the optimal gain can be described explicitly. Again, these algorithms are derived only for the systems with uncorrelated noises. One can also think about applying more complex estimation methods based on the cubature, unscented transformation or Monte-Carlo methods [12], [13], [14]. However, these methods are still too complex to be implemented on simple embedded devices which are commonly used nowadays. Also, even when such devices would have enough computational power, simple and optimal solutions are still going to be a better option.

The main contribution of this paper is the introduction of the filtering algorithms that take into account information

*This work has been supported by the project 18-26278S sponsored by Grant Agency of the Czech Republic and it has also been supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS16/232/OHK3/3T/13

[1]Department of Control Engineering, Faculty of Electrical Engineering of Czech Technical University in Prague, Czech Republic

978-1-5386-7891-6/18/$31.00 ©2018 IEEE

about the parameter uncertainty and noise correlation. The algorithms are based on the desensitized filtering method. First the general algorithm is proposed where the sensitivities are defined similarly to the sensitivities in the DKF. Then the special case of this algorithm is derived, where all formulas can be expressed explicitly. Also, the extended version for nonlinear systems is introduced. The performance of the filter is tested on the example, and it is compared to other estimation methods. The importance of the noise correlation information is demonstrated.

II. DESENSITIZED KALMAN FILTER ALGORITHMS FOR SYSTEMS WITH NOISE CORRELATION

This section introduces a derivation of the DKF algorithm for linear systems with uncertain parameters and correlated noises. A special case of this algorithm is also derived. The last subsection proposes an extended version for nonlinear systems.

A. Desensitized Kalman Filter for systems with Noise Correlation (DKF-NC)

Let's assume the linear uncertain stochastic system in the state space form

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \mathbf{A}(\boldsymbol{\rho})\boldsymbol{x}_k + \mathbf{B}(\boldsymbol{\rho})\mathbf{u}_k + \mathbf{G}(\boldsymbol{\rho})\mathbf{w}_k, \\ \mathbf{y}_k &= \mathbf{C}(\boldsymbol{\rho})\boldsymbol{x}_k + \mathbf{D}(\boldsymbol{\rho})\mathbf{u}_k + \mathbf{v}_k, \end{aligned} \tag{1}$$

where $\boldsymbol{x}$ is the state vector, $\boldsymbol{\rho}$ is the parameter vector, $\mathbf{u}$ is the input vector, $\mathbf{y}$ is the output vector and $\mathbf{w}$,$\mathbf{v}$ are the process and the measurement noise with the following statistics

$$\begin{bmatrix} \mathbf{w}_k \\ \mathbf{v}_k \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{Q} & \mathbf{S} \\ \mathbf{S}^T & \mathbf{R} \end{bmatrix} \right). \tag{2}$$

The notation that a matrix depends on the parameter is omitted in order to keep the text lucid (i.e. $\mathbf{A}(\boldsymbol{\rho}) \equiv \mathbf{A}$).

The single-step Kalman filter equations for correlated noises estimates the state vector optimally in terms of the minimization of the error covariance matrix. The update equations in Joseph's form are defined as follows

$$\hat{\boldsymbol{x}}_{k+1|k} = \mathbf{A}\hat{\boldsymbol{x}}_{k|k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{K}_k\left(\mathbf{y} - \mathbf{C}\hat{\boldsymbol{x}}_{k|k-1} - \mathbf{D}\mathbf{u}_k\right), \tag{3}$$

$$\begin{aligned} \mathbf{P}_{k+1|k} =& (\mathbf{A} - \mathbf{K}_k\mathbf{C})\mathbf{P}_{k|k-1}(\mathbf{A} - \mathbf{K}_k\mathbf{C})^T + \\ &+ \mathbf{Q} + \mathbf{K}_k\mathbf{R}\mathbf{K}_k^T - \mathbf{S}\mathbf{K}_k^T - \mathbf{K}_k\mathbf{S}^T, \end{aligned} \tag{4}$$

where $\hat{\boldsymbol{x}}$ represents the mean value of estimated state vector, $\mathbf{P}$ represents the state estimation error covariance, and $\mathbf{K}_k$ represents the optimal Kalman gain

$$\mathbf{K}_k = \left(\mathbf{A}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{S}\right)\left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{R}\right)^{-1}. \tag{5}$$

See Appendix I for the proof that the covariance update (4) keeps the covariance positive semi-definite. The DKF-NC is based on the updates (3) and (4) where $\mathbf{K}_k$ is derived from a modified optimality criterion, as it is shown in the next text.

Firstly, the state error sensitivity to the particular estimated parameter is defined as

$$\begin{aligned} \boldsymbol{\sigma}_{i,k+1|k} &= \frac{d\tilde{\boldsymbol{x}}_{k+1|k}}{d\rho_i} = \frac{d\left(\hat{\boldsymbol{x}}_{k+1|k} - \boldsymbol{x}_{k+1}\right)}{d\rho_i} = \\ &= \frac{d\hat{\boldsymbol{x}}_{k+1|k}}{d\rho_i} = \boldsymbol{\xi}_{i,k} - \mathbf{K}_k\boldsymbol{\gamma}_{i,k}, \end{aligned} \tag{6}$$

where the index i denotes the particular parameter in the parameter vector and

$$\boldsymbol{\xi}_{i,k} = \mathbf{A}\boldsymbol{\sigma}_{i,k|k-1} + \left(\partial\mathbf{A}/\partial\rho_i\right)\hat{\boldsymbol{x}}_{k|k-1} + \left(\partial\mathbf{B}/\partial\rho_i\right)\mathbf{u}_k, \tag{7}$$

$$\boldsymbol{\gamma}_{i,k} = \mathbf{C}\boldsymbol{\sigma}_{i,k|k-1} + \left(\partial\mathbf{C}/\partial\rho_i\right)\hat{\boldsymbol{x}}_{k|k-1} + \left(\partial\mathbf{D}/\partial\rho_i\right)\mathbf{u}_k. \tag{8}$$

Note that the true state is not sensitive to the expected value of the parameter which is set in the model description. To keep the derivation simple, it is assumed that $\partial\mathbf{K}_k/\partial\rho_i = 0$.

The optimality criterion for filtering problem is defined in the sense of desensitized filtering, which means there are two optimization objectives. The first objective is the minimization of the trace of the estimation error covariance matrix. The second is the minimization of the weighted state error sensitivities to the parameter. The criterion is defined as

$$J = \operatorname{tr}\left(\mathbf{P}_{k+1|k}\right) + \sum_{i=1}^{l}\left(\boldsymbol{\sigma}_{i,k+1|k}^T\mathbf{W}_i\boldsymbol{\sigma}_{i,k+1|k}\right), \tag{9}$$

where $\mathbf{W}_i$ are symmetric weighting matrices which represent the tuning parameter between two objectives in the optimality criterion. If $\mathbf{W}$ is large, then the emphasis is put on sensitivity, so the estimation is smoother. On the other hand, if it is set to a small value, then the accuracy is the main objective. If $\mathbf{W}$ is set to zero, then the algorithm becomes the conventional KF. To find the optimal gain $\mathbf{K}_k$, the partial derivative of the optimality criterion with respect to the gain

$$\begin{aligned} \frac{\partial J}{\partial\mathbf{K}_k} =& -2\left(\mathbf{A}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{S} + \sum_{i=1}^{l}\left(\mathbf{W}_i\boldsymbol{\xi}_{i,k}\boldsymbol{\gamma}_{i,k}^T\right)\right) + \\ &+ 2\mathbf{K}_k\left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{R}\right) + 2\sum_{i=1}^{l}\left(\mathbf{W}_i\mathbf{K}_k\boldsymbol{\gamma}_{i,k}\boldsymbol{\gamma}_{i,k}^T\right) \end{aligned} \tag{10}$$

are put equal to zero. Then the optimal gain is the solution of equation

$$\begin{aligned} &\mathbf{K}_k\left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{R}\right) + \sum_{i=1}^{l}\left(\mathbf{W}_i\mathbf{K}_k\boldsymbol{\gamma}_{i,k}\boldsymbol{\gamma}_{i,k}^T\right) = \\ &= \mathbf{A}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{S} + \sum_{i=1}^{l}\left(\mathbf{W}_i\boldsymbol{\xi}_{i,k}\boldsymbol{\gamma}_{i,k}^T\right). \end{aligned} \tag{11}$$

The derived algorithm consists of the update equations (3), (4) and (6) where $\mathbf{K}_k$ is the solution of (11). The summary of the algorithm is in Fig. 1.

B. Special case of DKF-NC (SDKF-NC)

The DKF-NC algorithm requires solving the implicit equation at each update step which creates a significant computational burden. One of the reasons is that the definition of sensitivity is too general. Let's create a special case where the sensitivity is defined as the state error vector sensitivity to the parameter vector

$$\begin{aligned} \boldsymbol{\Sigma}_{k+1|k} &= \frac{d\tilde{\boldsymbol{x}}_{k+1|k}}{d\boldsymbol{\rho}} = \frac{d\left(\hat{\boldsymbol{x}}_{k+1|k} - \boldsymbol{x}_{k+1}\right)}{d\boldsymbol{\rho}} = \\ &= \frac{d\hat{\boldsymbol{x}}_{k+1|k}}{d\boldsymbol{\rho}} = \boldsymbol{\Xi}_k - \mathbf{K}_k\boldsymbol{\Gamma}_k, \end{aligned} \tag{12}$$

$$
\begin{aligned}
&\boldsymbol{\xi}_{i,k} = \mathbf{A}\boldsymbol{\sigma}_{i,k|k-1} + \left(\partial \mathbf{A}/\partial \rho_i\right)\hat{\boldsymbol{x}}_{k|k-1} + \left(\partial \mathbf{B}/\partial \rho_i\right)\mathbf{u}_k,\\
&\boldsymbol{\gamma}_{i,k} = \mathbf{C}\boldsymbol{\sigma}_{i,k|k-1} + \left(\partial \mathbf{C}/\partial \rho_i\right)\hat{\boldsymbol{x}}_{k|k-1} + \left(\partial \mathbf{D}/\partial \rho_i\right)\mathbf{u}_k,\\
&\text{Solve for } \mathbf{K}_k:\\
&\mathbf{K}_k\left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{R}\right) + \sum_{i=1}^{l}\left(\mathbf{W}_i\mathbf{K}_k\boldsymbol{\gamma}_{i,k}\boldsymbol{\gamma}_{i,k}^T\right) =\\
&\qquad = \mathbf{A}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{S} + \sum_{i=1}^{l}\left(\mathbf{W}_i\boldsymbol{\xi}_{i,k}\boldsymbol{\gamma}_{i,k}^T\right),\\
\hline
&\boldsymbol{\sigma}_{i,k+1|k} = \boldsymbol{\xi}_{i,k} - \mathbf{K}_k\boldsymbol{\gamma}_{i,k},\\
&\hat{\boldsymbol{x}}_{k+1|k} = \mathbf{A}\hat{\boldsymbol{x}}_{k|k-1} + \mathbf{B}\mathbf{u}_k +\\
&\qquad + \mathbf{K}_k\left(\mathbf{y} - \mathbf{C}\hat{\boldsymbol{x}}_{k|k-1} - \mathbf{D}\mathbf{u}_k\right),\\
&\mathbf{P}_{k+1|k} = (\mathbf{A} - \mathbf{K}_k\mathbf{C})\mathbf{P}_{k|k-1}(\mathbf{A} - \mathbf{K}_k\mathbf{C})^T +\\
&\qquad + \mathbf{Q} + \mathbf{K}_k\mathbf{R}\mathbf{K}_k^T - \mathbf{S}\mathbf{K}_k^T - \mathbf{K}_k\mathbf{S}^T.
\end{aligned}
$$

Fig. 1. DKF-NC algorithm: DKF for systems with noise correlation

where

$$
\begin{aligned}
\boldsymbol{\Xi}_k &= \mathbf{A}\boldsymbol{\Sigma}_{k|k-1} + \left(\partial \mathbf{A}/\partial \boldsymbol{\rho}\right)\hat{\boldsymbol{x}}_{k|k-1} + \left(\partial \mathbf{B}/\partial \boldsymbol{\rho}\right)\mathbf{u}_k, && (13)\\
\boldsymbol{\Gamma}_k &= \mathbf{C}\boldsymbol{\Sigma}_{k|k-1} + \left(\partial \mathbf{C}/\partial \boldsymbol{\rho}\right)\hat{\boldsymbol{x}}_{k|k-1} + \left(\partial \mathbf{D}/\partial \boldsymbol{\rho}\right)\mathbf{u}_k. && (14)
\end{aligned}
$$

Again, the derivation is simplified by using the assumption $\partial \mathbf{K}_k/\partial \boldsymbol{\rho} = 0$. Then the second objective of optimality criterion changes to the minimization of the trace of the weighted sensitivity matrix, so now the optimality criterion is defined as

$$
J = \operatorname{tr}\left(\mathbf{P}_{k+1|k}\right) + \operatorname{tr}\left(\boldsymbol{\Sigma}_{k+1|k}\mathbf{W}\boldsymbol{\Sigma}_{k+1|k}^T\right). \tag{15}
$$

Note the weighting matrices in (9) are the symmetric matrices $\mathbf{W}_i \in \mathbb{R}^{n\times n}$, where n is the size of state vector and the weighting matrix in (15) is the symmetric matrix $\mathbf{W} \in \mathbb{R}^{p\times p}$, where p is the size of parameter vector. The weighting in the DKF-NC can be tuned more accurately since it is possible to weight each parameter for each state separately. The weights in the SDKF-NC set same weights of parameter vector for all states.

The partial derivative of the criterion with respect to the gain is obtained as

$$
\begin{aligned}
\frac{\partial J}{\partial \mathbf{K}_k} =& -2\left(\mathbf{A}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{S} + \boldsymbol{\Xi}_k\mathbf{W}\boldsymbol{\Gamma}_k^T\right) +\\
&+ 2\mathbf{K}_k\left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{R} + \boldsymbol{\Gamma}_k\mathbf{W}\boldsymbol{\Gamma}_k^T\right).
\end{aligned} \tag{16}
$$

The optimal gain is the solution of the equation, which is created by putting the above derivative equal to zero. Then the optimal gain is obtained by the explicit formula

$$
\begin{aligned}
\mathbf{K}_k =& \left(\mathbf{A}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{S} + \boldsymbol{\Xi}_k\mathbf{W}\boldsymbol{\Gamma}_k^T\right) \times\\
&\times \left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{R} + \boldsymbol{\Gamma}_k\mathbf{W}\boldsymbol{\Gamma}_k^T\right)^{-1}.
\end{aligned} \tag{17}
$$

The covariance update equation can be further modified by substituting the optimal gain defined in (17) into (4). After the substitution, the alternative form of the covariance update equation is obtained as follows

$$
\begin{aligned}
\mathbf{P}_{k+1|k} =& \mathbf{A}\mathbf{P}_{k|k-1}\mathbf{A}^T + \mathbf{Q} -\\
&- \mathbf{K}_k\left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{A}^T + \mathbf{S}^T\right) + \boldsymbol{\Sigma}_{k+1|k}\mathbf{W}\boldsymbol{\Gamma}_k^T\mathbf{K}_k^T.
\end{aligned} \tag{18}
$$

To clarify that the alternative covariance update results in a symmetric matrix, the formulas (12) and (17) can be substituted in (18). Then the obtained formulation is clearly a symmetric matrix

$$
\begin{aligned}
\mathbf{P}_{k+1|k} =& \mathbf{A}\mathbf{P}_{k|k-1}\mathbf{A}^T + \mathbf{Q}\\
&- \left(\mathbf{A}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{S}\right)\mathbf{P}_y^{-1}\left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{A}^T + \mathbf{S}^T\right)\\
&- \mathbf{K}_k\boldsymbol{\Gamma}_k\mathbf{W}\boldsymbol{\Gamma}_k^T\mathbf{K}_k^T + \left(\boldsymbol{\Xi}_k\mathbf{W}\boldsymbol{\Gamma}_k^T\right)\mathbf{P}_y^{-1}\left(\boldsymbol{\Xi}_k\mathbf{W}\boldsymbol{\Gamma}_k^T\right)^T,
\end{aligned} \tag{19}
$$

where

$$
\mathbf{P}_y = \mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{R} + \boldsymbol{\Gamma}_k\mathbf{W}\boldsymbol{\Gamma}_k^T. \tag{20}
$$

The special case of DKF-NC (SDKF-NC) algorithm consists of the update equations (3),(4) (or (18)) and (12) where $\mathbf{K}_k$ is defined in (17). This algorithm updates the statistics with the explicit formulas, therefore, it is more efficient than the DKF-NC. The algorithm is summed up in Fig. 2.

$$
\begin{aligned}
\boldsymbol{\Xi}_k &= \mathbf{A}\boldsymbol{\Sigma}_{k|k-1} + \left(\partial \mathbf{A}/\partial \boldsymbol{\rho}\right)\hat{\boldsymbol{x}}_{k|k-1} + \left(\partial \mathbf{B}/\partial \boldsymbol{\rho}\right)\mathbf{u}_k,\\
\boldsymbol{\Gamma}_k &= \mathbf{C}\boldsymbol{\Sigma}_{k|k-1} + \left(\partial \mathbf{C}/\partial \boldsymbol{\rho}\right)\hat{\boldsymbol{x}}_{k|k-1} + \left(\partial \mathbf{D}/\partial \boldsymbol{\rho}\right)\mathbf{u}_k,\\
\mathbf{K}_k &= \left(\mathbf{A}_{x,k}\mathbf{P}_{k|k-1}\mathbf{C}_{x,k}^T + \mathbf{S} + \boldsymbol{\Xi}_k\mathbf{W}\boldsymbol{\Gamma}_k^T\right) \times\\
&\quad \times \left(\mathbf{C}_{x,k}\mathbf{P}_{k|k-1}\mathbf{C}_{x,k}^T + \mathbf{R} + \boldsymbol{\Gamma}_k\mathbf{W}\boldsymbol{\Gamma}_k^T\right)^{-1},\\
\hline
\boldsymbol{\Sigma}_{k+1|k} &= \boldsymbol{\Xi}_k - \mathbf{K}_k\boldsymbol{\Gamma}_k,\\
\hat{\boldsymbol{x}}_{k+1|k} &= \mathbf{A}\hat{\boldsymbol{x}}_{k|k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{K}_k\left(\mathbf{y} - \mathbf{C}\hat{\boldsymbol{x}}_{k|k-1} - \mathbf{D}\mathbf{u}_k\right),\\
\mathbf{P}_{k+1|k} &= (\mathbf{A} - \mathbf{K}_k\mathbf{C})\mathbf{P}_{k|k-1}(\mathbf{A} - \mathbf{K}_k\mathbf{C})^T +\\
&\quad + \mathbf{Q} + \mathbf{K}_k\mathbf{R}\mathbf{K}_k^T - \mathbf{S}\mathbf{K}_k^T - \mathbf{K}_k\mathbf{S}^T.\\
&\text{or}\\
\mathbf{P}_{k+1|k} &= \mathbf{A}\mathbf{P}_{k|k-1}\mathbf{A}^T + \mathbf{Q} -\\
&\quad - \mathbf{K}_k\left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{A}^T + \mathbf{S}^T\right) + \boldsymbol{\Sigma}_{k+1|k}\mathbf{W}\boldsymbol{\Gamma}_k^T\mathbf{K}_k^T
\end{aligned}
$$

Fig. 2. SDKF-NC algorithm: Special DKF for systems with noise correlation

C. Extended SDKF-NC (ESDKF-NC)

The extension for nonlinear systems is derived similarly. Note that only the extended version of SDKF-NC is shown because this algorithm is more useful in practical applications. The extended version of DKF-NC can be derived similarly.

Let's assume the nonlinear uncertain stochastic system in the following state space form

$$
\begin{aligned}
\boldsymbol{x}_{k+1} &= \mathbf{f}(\boldsymbol{x}_k, \mathbf{u}_k, \boldsymbol{\rho}) + \mathbf{w}_k,\\
\mathbf{y}_k &= \mathbf{h}(\boldsymbol{x}_k, \mathbf{u}_k, \boldsymbol{\rho}) + \mathbf{v}_k,
\end{aligned} \tag{21}
$$

where $\mathbf{f}$, $\mathbf{h}$ are nonlinear functions of state update and output measurements respectively. The Jacobians of these functions are defined as

$$\begin{aligned}\mathbf{A}_{x,k} &= \left.\frac{\partial \mathbf{f}}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k|k-1},\mathbf{u}_k,\boldsymbol{\rho}}, \mathbf{A}_{\rho,k} = \left.\frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}}\right|_{\hat{\boldsymbol{x}}_{k|k-1},\mathbf{u}_k,\boldsymbol{\rho}}, \\ \mathbf{C}_{x,k} &= \left.\frac{\partial \mathbf{h}}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k|k-1},\mathbf{u}_k,\boldsymbol{\rho}}, \mathbf{C}_{\rho,k} = \left.\frac{\partial \mathbf{h}}{\partial \boldsymbol{\rho}}\right|_{\hat{\boldsymbol{x}}_{k|k-1},\mathbf{u}_k,\boldsymbol{\rho}}.\end{aligned} \tag{22}$$

The update equations of the single-step extended Kalman filter in Joseph's form are as follows

$$\hat{\boldsymbol{x}}_{k+1|k} = \mathbf{f}\left(\hat{\boldsymbol{x}}_{k|k-1}, \mathbf{u}_k, \boldsymbol{\rho}\right) + \mathbf{K}_k\left(\mathbf{y} - \mathbf{h}\left(\hat{\boldsymbol{x}}_{k|k-1}, \mathbf{u}_k, \boldsymbol{\rho}\right)\right), \tag{23}$$

$$\begin{aligned}\mathbf{P}_{k+1|k} =& (\mathbf{A}_{x,k} - \mathbf{K}_k\mathbf{C}_{x,k})\mathbf{P}_{k|k-1}(\mathbf{A}_{x,k} - \mathbf{K}_k\mathbf{C}_{x,k})^T + \\ &+ \mathbf{Q} + \mathbf{K}_k\mathbf{R}\mathbf{K}_k^T - \mathbf{S}\mathbf{K}_k^T - \mathbf{K}_k\mathbf{S}^T,\end{aligned} \tag{24}$$

where $\mathbf{K}_k$ is the optimal Kalman gain which will be rederived using the new optimality criterion. The state error vector sensitivity to the parameter vector is defined similarly to (12). It is defined in the form of the sensitivity matrix

$$\boldsymbol{\Sigma}_{k+1|k} = \frac{d\hat{\boldsymbol{x}}_{k+1|k}}{d\boldsymbol{\rho}} = \boldsymbol{\Xi}_k - \mathbf{K}_k\boldsymbol{\Gamma}_k, \tag{25}$$

where

$$\boldsymbol{\Xi}_k = \mathbf{A}_{x,k}\boldsymbol{\Sigma}_{k|k-1} + \mathbf{A}_{\rho,k}, \tag{26}$$

$$\boldsymbol{\Gamma}_k = \mathbf{C}_{x,k}\boldsymbol{\Sigma}_{k|k-1} + \mathbf{C}_{\rho,k}. \tag{27}$$

The optimality criterion is the same as the criterion in the SKDF-NC, which is

$$J = \operatorname{tr}\left(\mathbf{P}_{k+1|k}\right) + \operatorname{tr}\left(\boldsymbol{\Sigma}_{k+1|k}\mathbf{W}\boldsymbol{\Sigma}_{k+1|k}^T\right). \tag{28}$$

The derivation of the optimal gain is also identical to the SDKF-NC. The optimal gain which results from the minimization of (28) is

$$\begin{aligned}\mathbf{K}_k =& \left(\mathbf{A}_{x,k}\mathbf{P}_{k|k-1}\mathbf{C}_{x,k}^T + \mathbf{S} + \boldsymbol{\Xi}_k\mathbf{W}\boldsymbol{\Gamma}_k^T\right) \times \\ &\times \left(\mathbf{C}_{x,k}\mathbf{P}_{k|k-1}\mathbf{C}_{x,k}^T + \mathbf{R} + \boldsymbol{\Gamma}_k\mathbf{W}\boldsymbol{\Gamma}_k^T\right)^{-1}.\end{aligned} \tag{29}$$

The optimal gain (29) can be substituted into the covariance update equation (24), which will again give the alternative covariance update (18).

The Extended SDKF-NC (ESDKF-NC) algorithm consists of the update equations (23), (24) (or (18)) and (25) where $\mathbf{K}_k$ is defined in (29). The algorithm is summed up in Fig. 3.

$$\begin{aligned}\mathbf{A}_{x,k} &= \left.\frac{\partial \mathbf{f}}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k|k-1},\mathbf{u}_k,\boldsymbol{\rho}}, \mathbf{A}_{\rho,k} = \left.\frac{\partial \mathbf{f}}{\partial \boldsymbol{\rho}}\right|_{\hat{\boldsymbol{x}}_{k|k-1},\mathbf{u}_k,\boldsymbol{\rho}}, \\ \mathbf{C}_{x,k} &= \left.\frac{\partial \mathbf{h}}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k|k-1},\mathbf{u}_k,\boldsymbol{\rho}}, \mathbf{C}_{\rho,k} = \left.\frac{\partial \mathbf{h}}{\partial \boldsymbol{\rho}}\right|_{\hat{\boldsymbol{x}}_{k|k-1},\mathbf{u}_k,\boldsymbol{\rho}}, \\ \boldsymbol{\Xi}_k &= \mathbf{A}_{x,k}\boldsymbol{\Sigma}_{k|k-1} + \mathbf{A}_{\rho,k}, \\ \boldsymbol{\Gamma}_k &= \mathbf{C}_{x,k}\boldsymbol{\Sigma}_{k|k-1} + \mathbf{C}_{\rho,k}, \\ \mathbf{K}_k &= \left(\mathbf{A}_{x,k}\mathbf{P}_{k|k-1}\mathbf{C}_{x,k}^T + \mathbf{S} + \boldsymbol{\Xi}_k\mathbf{W}\boldsymbol{\Gamma}_k^T\right) \times \\ &\times \left(\mathbf{C}_{x,k}\mathbf{P}_{k|k-1}\mathbf{C}_{x,k}^T + \mathbf{R} + \boldsymbol{\Gamma}_k\mathbf{W}\boldsymbol{\Gamma}_k^T\right)^{-1},\end{aligned}$$

$$\begin{aligned}\boldsymbol{\Sigma}_{k+1|k} =& \boldsymbol{\Xi}_k - \mathbf{K}_k\boldsymbol{\Gamma}_k, \\ \hat{\boldsymbol{x}}_{k+1|k} =& \mathbf{f}\left(\hat{\boldsymbol{x}}_{k|k-1}, \mathbf{u}_k, \boldsymbol{\rho}\right) + \mathbf{K}_k\left(\mathbf{y} - \mathbf{h}\left(\hat{\boldsymbol{x}}_{k|k-1}, \mathbf{u}_k, \boldsymbol{\rho}\right)\right), \\ \mathbf{P}_{k+1|k} =& (\mathbf{A}_{x,k} - \mathbf{K}_k\mathbf{C}_{x,k})\mathbf{P}_{k|k-1}(\mathbf{A}_{x,k} - \mathbf{K}_k\mathbf{C}_{x,k})^T + \\ &+ \mathbf{Q} + \mathbf{K}_k\mathbf{R}\mathbf{K}_k^T - \mathbf{S}\mathbf{K}_k^T - \mathbf{K}_k\mathbf{S}^T. \\ &\text{or} \\ \mathbf{P}_{k+1|k} =& \mathbf{A}_{x,k}\mathbf{P}_{k|k-1}\mathbf{A}_{x,k}^T + \mathbf{Q} - \\ &- \mathbf{K}_k\left(\mathbf{C}_{x,k}\mathbf{P}_{k|k-1}\mathbf{A}_{x,k}^T + \mathbf{S}^T\right) + \boldsymbol{\Sigma}_{k+1|k}\mathbf{W}\boldsymbol{\Gamma}_k^T\mathbf{K}_k^T\end{aligned}$$

Fig. 3. ESDKF-NC algorithm: Extended Special DKF for systems with Noise Correlation

III. PERFORMANCE

The importance of introduced algorithms is discussed in this section. The SDKF-NC was chosen as a baseline for the algorithms proposed in this paper. The first reason is that the SDKF-NC is a more efficient algorithm than the general DKF-NC and it is also more suitable for real applications. Secondly, the performance of the extended version (ESDKF-NC) is unsuitable for the performance comparison because a linearization error could influence the results.

The performance is tested against the following state estimation algorithms. The first is the "Perfect" Kalman Filter (PKF) which represents the theoretical case which has the information about true parameter value. Such Kalman filter reaches the Cramer-Rao bound, therefore, it is the best estimator. The second tested algorithm is the "Imperfect" Kalman Filter (IKF), which is the Kalman filter where the parameter is set to the expected value. In other words, it is the Kalman filter with a biased model of system. The last compared algorithm is the SDKF introduced in [11] which is the special case of DKF algorithm that assumes uncorrelated noises.

The example for the performance testing is chosen to be a continuous-time system with discrete-time measurements. The reason for selecting such a system is that firstly, it is more natural to describe physical phenomena in the continuous-time domain. Secondly, system outputs are measured by devices in the discrete-time domain. Hence the discrete-time measurement model is an obvious choice. The continuous-discrete state estimators are available but it is more practical to use the discrete state estimation, where a continuous system is discretized first.

A. Discretization of continuous system

The same models are usually used for estimation and control at the same time. However, often the measurement and control input sampling is offset. This method is called asynchronous sampling, and it is used to pass the maximum amount of the information to the control algorithm. The asynchronous sampling method is depicted in Fig. 4. The control input sampling period T_s is split by the parameter ε. The period $T_c = (1-\varepsilon)T_s$ represents the time needed for computation of the new input, and it is usually minimized in order to supply the control algorithm with the latest information. Such sampling configuration causes correlation between process and measurement noise which occurs in period εT_s. Such scenario is a typical example how a continuous-discrete system becomes a system with noise correlation just because of the sampling method.

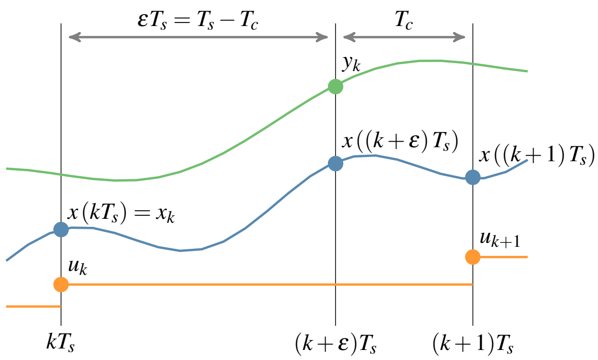

Fig. 4. Asynchronous sampling of a system with control.

B. Example

The particular example for performance tests is a simple first-order continuous system with one uncertain parameter. The continuous-discrete system description is

$$dx(t) = \frac{1}{\rho}\left(3u(t) - x(t)\right)dt + dw(t), \tag{30}$$

$$y(t) = 100x(t) + v(t), \tag{31}$$

where ρ is the uncertain parameter which represents the uncertain time constant of the first-order system. The system is discretized in sampling period $T_s = 0.5$ s, and the measurements are asynchronous with $\varepsilon = 0.75$. The parameter is expected to be from interval $\rho \in [0.5, 7.5]$, with $\hat{\rho} = 4$ used as the parameter value in the model used in the SDKF, SKDF-NC and IKF algorithms. The performance test consists of the system and state estimators response to the single pulse with amplitude 1 and length 30 s. The SDKF-NC and SDKF algorithms are tested for multiple weight settings from the interval $\left[10^{-1}, 10^{3}\right]$ to try different trade-offs between two optimization objectives. Each test configuration is repeated in 100 Monte Carlo simulations, where the true value of the parameter is generated randomly from the defined interval of parameter values. The performance was tested on systems with various process and measurement noise covariances. However, the estimation algorithms always used the true (optimal) values of noise covariances.

C. Results

The results are shown in the form of evolution of the mean NRMSE (Normalized Root Mean Square Error) in time for particular noise covariances setting.

The scenario with the high process noise is shown in Fig. 5. It can be seen that the SDKF-NC is almost an efficient estimator when the optimal weight is set. The IKF is also close to the PKF since the measurement noise has small variance, and therefore KF trusts the measurements more than the model. SDKF gives the worst performance in this scenario, and even the SDKF with the best weighting underperforms the SDKF-NC with the worst weighting.

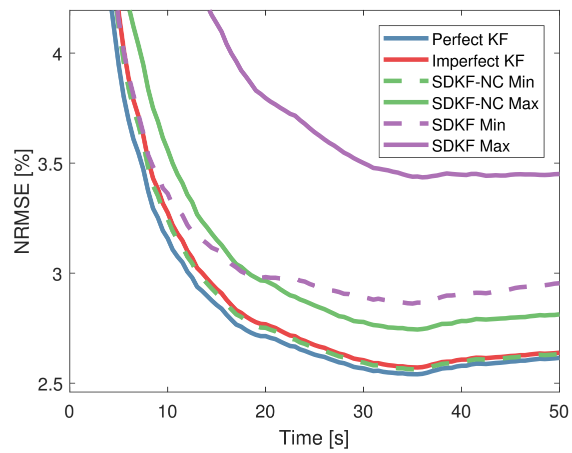

Fig. 5. Mean NRMSE (from 100 Monte Carlo simulations) for scenario with noise statistics: $Q = 1, R = 0.001$. Desensitized algorithms are shown for weights settings which achieved maximal and minimal mean NRMSE ($W_{SDKFNC,Min} = 10, W_{SDKFNC,Max} = 10^3, W_{SDKF,Min} = 10^2, W_{SDKF,Max} = 10^{-1}$)

The situation depicted in Fig. 6 shows the configuration where both noises have small covariances. The performance results are very similar to the previous scenario. It can be seen that the SDKF-NC is closer to the PKF than any other estimator. The weight setting does not have a substantial impact on the accuracy, which means that less tuning is required. The SDKF performs even worse than the IKF, but despite that, the steady-state error is only ≈ 1 % for the SDKF with the best setting. On the other hand, the accuracy moves within 0.5 % depending on the weighting, so more tuning is required to get the best result.

The scenario shown in Fig. 7 represents the case when the process noise covariance is large, and the measurement covariance is small. In this case, the correlation information is less significant due to the high measurement noise. The best settings of the SDKF and SDKF-NC give performance close to the efficiency bound. However, the weighting of the SDKF must be adequately tuned, since the worst weighting gives worse results than the IKF. On the other hand, the SDKF-NC performs close to the efficiency bound for all weight settings.

The results show that no single weight is the best for all scenarios. The optimal weight setting depends on the particular application and it needs to be tuned using simulations.

IV. CONCLUSION

This paper generalized an approach to the desensitized optimal filtering by including the information about noise correlation. The general DKF-NC algorithm was derived, but its computational complexity could restrain its potential for real application. Therefore, the special case of DKF-NC

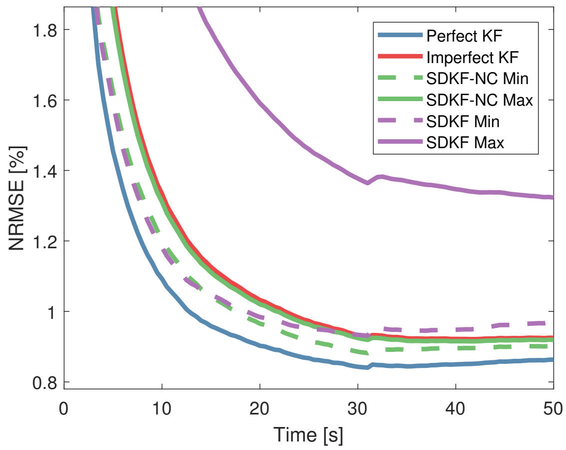

Fig. 6. Mean NRMSE (from 100 Monte Carlo simulations) for scenario with noise statistics: $Q = 0.1, R = 0.1$. Desensitized algorithms are shown for weights settings which achieved maximal and minimal mean NRMSE ($W_{SDKFNC,Min} = 10^2, W_{SDKFNC,Max} = 10^{-1}, W_{SDKF,Min} = 10^2, W_{SDKF,Max} = 10^{-1}$)

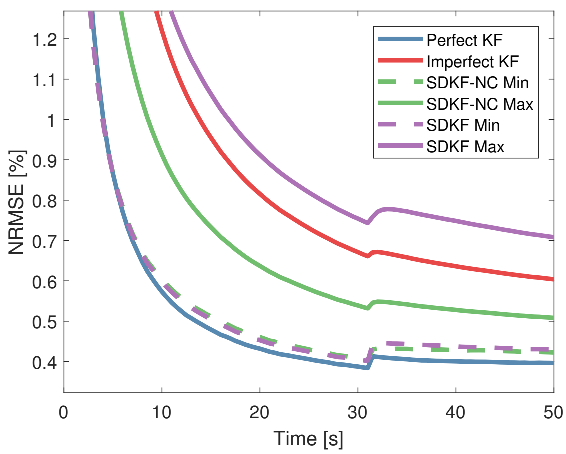

Fig. 7. Mean NRMSE (from 100 Monte Carlo simulations) for scenario with noise statistics: $Q = 0.01, R = 10$. Desensitized algorithms are shown for weights settings which achieved maximal and minimal mean NRMSE ($W_{SDKFNC,Min} = 10^2, W_{SDKFNC,Max} = 10^{-1}, W_{SDKF,Min} = 10^2, W_{SDKF,Max} = 10^{-1}$)

(SDKF-NC) was shown which has low complexity and it is more suitable for practical applications. Also, the extended version for nonlinear systems was proposed. All of these algorithms can also be used for systems without noise correlation. Moreover, the weighting of the sensitivities in the algorithm allows tuning the trade-off between the overall accuracy and the smoothness of state estimation. The performance of the algorithm was tested on the example where the information about noise correlation is not negligible. The importance of including this information was shown to be beneficial since the newly proposed algorithm with the noise correlation information (SDKF-NC) outperformed the algorithm which assumes uncorrelated noises (SDKF).

Appendix I
Proof of positive semi-definiteness of error covariance matrix

The joint probability distribution of noises is described in (2). A covariance matrix is always positive semi-definite. Therefore, the covariance of noises distribution (2) is also positive semi-definite

$$\begin{bmatrix} \mathbf{Q} & \mathbf{S} \\ \mathbf{S}^T & \mathbf{R} \end{bmatrix} \geq 0. \tag{32}$$

Then the following matrix is also positive semi-definite [15]

$$\begin{aligned} \begin{bmatrix} \mathbf{I}_x & -\mathbf{K}_k \end{bmatrix} \begin{bmatrix} \mathbf{Q} & \mathbf{S} \\ \mathbf{S}^T & \mathbf{R} \end{bmatrix} \begin{bmatrix} \mathbf{I}_x \\ -\mathbf{K}_k^T \end{bmatrix} = \\ = \mathbf{Q} - \mathbf{K}_k\mathbf{S}^T - \mathbf{S}\mathbf{K}_k^T + \mathbf{K}_k\mathbf{R}\mathbf{K}_k^T \geq 0, \end{aligned} \tag{33}$$

where $\mathbf{I}_x$ is the identity matrix of size n_x and n_x denotes the number of elements in the state vector. The covariance update is defined in (4). It is defined as the sum of the positive semi-definite matrix in (33) and the matrix $(\mathbf{A} - \mathbf{K}_k\mathbf{C})\mathbf{P}_{k|k-1}(\mathbf{A} - \mathbf{K}_k\mathbf{C})^T$ which is positive semi-definite. Then the entire update is the sum of positive semi-definite matrices which results in a positive semi-definite matrix. ■

References

[1] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Burlington: Academic Press, 1970, vol. 64.

[2] S. F. Schmidt, “Applications of state space methods to navigation problems,” *in C. T. Leondes, Editor, Advanced Control Systems*, vol. 3, pp. 293–340, 1966.

[3] D. Woodbury and J. Junkins, “On the consider kalman filter,” in *AIAA Guidance, Navigation, and Control Conference*, Toronto, Ontario, Canada, 2010.

[4] D. Woodbury, “Accounting for parameter uncertainty in reduced-order static and dynamic systems,” Ph.D. dissertation, Texas A&M University, 2011.

[5] J. Chen, R. J. Patton, and H.-Y. Zhang, “Design of unknown input observers and robust fault detection filters,” *International Journal of Control*, vol. 63, no. 1, pp. 85–105, 1996.

[6] M. Corless and J. Tu, “State and input estimation for a class of uncertain systems,” *Automatica*, vol. 34, no. 6, pp. 757–764, 1998.

[7] D. Simon, *Optimal state estimation: Kalman, H infinity, and Nonlinear Approaches*. Hoboken: Wiley, 2006.

[8] Y. Hung and F. Yang, “Robust H∞ filtering with error variance constraints for discrete time-varying systems with uncertainty,” *Automatica*, vol. 39, no. 7, pp. 1185–1194, 2003.

[9] L. Xie, Y. C. Soh, and C. de Souza, “Robust kalman filtering for uncertain discrete-time systems,” *IEEE Transactions on Automatic Control*, vol. 39, no. 6, pp. 1310–1314, 1994.

[10] C. D. Karlgaard and H. Shen, “Desensitised kalman filtering,” *IET Radar, Sonar & Navigation*, vol. 7, no. 1, pp. 2–9, 2013.

[11] T. Lou, “Desensitized kalman filtering with analytical gain,” *CoRR*, vol. abs/1504.04916, 2015. [Online]. Available: http://arxiv.org/abs/1504.04916

[12] T. Lou, L. Wang, H. Su, M.-W. Nie, N. Yang, and Y. Wang, “Desensitized cubature kalman filter with uncertain parameters,” *Journal of the Franklin Institute*, vol. 354, no. 18, pp. 8358–8373, 2017.

[13] H. Shen and C. D. Karlgaard, “Sensitivity reduction of unscented kalman filter about parameter uncertainties,” *IET Radar, Sonar & Navigation*, vol. 9, no. 4, pp. 374–383, 2015.

[14] S. Ishihara and M. Yamakita, “Robust approximated unbiased minimum variance filter for nonlinear systems with parameter uncertainties,” in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017.

[15] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge: Cambridge University Press, 2013.