Adaptive Sigma-Point Kalman Filtering for Wind Turbine State and Process Noise Estimation

B Ritter\textsuperscript{1,2}, E Mora\textsuperscript{1}, T Schlicht\textsuperscript{2}, A Schild\textsuperscript{3}, U Konigorski\textsuperscript{1}

\textsuperscript{1} Technische Universität Darmstadt, Department of Control Systems and Mechatronics, DE
\textsuperscript{2} Industrial Science GmbH, Alexanderstr. 25, 64283 Darmstadt, DE
\textsuperscript{3} IAV GmbH, Rockwellstraße 16, 38518 Gifhorn, DE

E-mail: britter@iat.tu-darmstadt.de

Abstract. This contribution investigates different noise adaptive Kalman filtering techniques with regard to their usability for wind turbine application. Since advanced model-based control schemes arise as promising alternative for standard industrial control, the necessity for robust and adaptive state estimation techniques has simultaneously emerged as an important topic. The comparison of the implemented adaptation rules shows that the master-slave adaptive filters are very flexible, numerically efficient and easy to implement. Maximum likelihood estimation based methods are more robust but show less flexibility and fewer design parameters to influence the filter performance. The simulation study shows that adaptive filters are beneficial since they solve two typical problems involved with static Kalman filter design: First, filter parameter adaptation compensates incorrect assumptions of noise statistics. Secondly, adaptation rules prevent poor filter performance for systems with time-varying statistical properties.

1. Introduction
The precise knowledge about the current dynamic state including nacelle, drive-train and blade dynamics is vital information for advanced wind turbine control schemes [1, 2, 3, 4]. Since load sensors like optic strain gauges are often cost-intensive, error-prone and with limited technical availability, state estimation algorithms as inexpensive software add-on are preferable either for reconstructing unmeasured quantities using data fusion methods. Previous work has focused in detail on the full-scope estimation problem for wind turbines and how to tackle it with a distributed observer architecture [5]. Moreover, real-time capabilities, state observability and benefits of additional advanced measurement configurations have been discussed in [6]. In the presence of stochastic disturbances, Kalman Filters (KF) are a convenient choice. The inherent issues of the extended Kalman filter as the standard approach applied to nonlinear systems can be resolved since the invention of the sigma-point Kalman filter (SPKF). However, the problem of finding the suitable noise covariance matrices continues to exist. Since the latter strongly affect the filter performance, the practitioner is concerned with selecting good noise parameters to ensure the expected estimation accuracy. Numerical adaptation methods can support, improve and accelerate this design process. The remainder of this paper is structured as follows: Sec. 2 introduces the basics about nonlinear adaptive Kalman filters. Sec. 3 elaborates on the internal filter model and the observability assessment. Sec. 4 elaborates on the pursued design approach. Sec. 5 defines the estimation problems and discusses the simulation results. Finally, Sec. 6 summarizes the key statements and provides a glimpse on the proposed future work.
2. Noise adaptive Kalman filters
This section introduces briefly the concepts of nonlinear sigma-point Kalman filters (SPKF) and adaptive Kalman filtering for process noise estimation. The details on linear Kalman filters (LKF) for linear Gaussian estimation problems are presented in [7, 8, 9]. A nomenclature for the used symbols is provided in Tab. 1.

2.1. Sigma-point Kalman filters
For a long time the most important algorithm in industrial practice for nonlinear estimation problems has been the Extended Kalman Filter (EKF) despite some serious drawbacks [10]. Fortunately, the ideas of JULIER and UHLMANN have provoked the invention of the unscented Kalman filter (UKF) [11, 12]. The SPKF is an estimation algorithm for dynamic state-space systems governed by a set of nonlinear ordinary differential equations

\[ x_{k+1} = f(x_k, u_k, d_k, \theta_k) + w_k \quad \text{with} \quad w_k \sim \mathcal{N}(0, Q_k) \]  
\[ y_k = h(x_k, u_k, d_k, \theta_k) + n_k \quad \text{with} \quad n_k \sim \mathcal{N}(0, R_k) \]  

(1a) 

(1b)

The cubature Kalman filter (CKF) is used in the following as representative type of SPKF. For sake of completeness, its equations are presented here in the most economical way:

\[ \mathbf{X}^+_k = \tilde{\mathbf{X}}^+_k + \sqrt{n} \left[ \sqrt{\mathbf{P}^+_k}, -\sqrt{\mathbf{P}^+_k} \right] \]  
\[ \mathbf{X}^-_k = \mathbf{f}(\mathbf{X}^+_k, \mathbf{U}_k) \]  
\[ \bar{x}_k = \frac{1}{2n} \sum_{i=0}^{2n} \mathbf{X}^+_k(:, i) \]  
\[ \mathbf{P}^-_k = \frac{1}{2n} \mathbf{X}^+_k(\mathbf{X}^+_k)^T - \bar{x}_k(\bar{x}_k)^T + Q_{k-1} \]  
\[ \mathbf{Y}^+_k = \tilde{\mathbf{X}}^-_k + \sqrt{n} \left[ \sqrt{\mathbf{P}^-_k}, -\sqrt{\mathbf{P}^-_k} \right] \]  
\[ \mathbf{Y}^-_k = \mathbf{h}(\mathbf{X}^-_k, \mathbf{U}_k) \]  
\[ \hat{y}_k = \frac{1}{2n} \sum_{i=0}^{2n} \mathbf{Y}^+_k(:, i) \]  

A more detailed review and the derivations can be found in [13].

2.2. Process and measurement noise adaptation
Although the standard SPKF (see Fig.1) tackles the problem of system nonlinearity effectively, it cannot address the problem of changing and/or unknown noise statistics by nature. Lack of knowledge about these noise statistics is often present in technical systems. Unfortunate design may lead to filter divergence or severe performance degradation in practice. One way to deal with this issue is a very conservative design of the noise matrices \( Q_k \) and \( R_k \).

A more elegant approach is to set up an estimation problem (EP) and solve it numerically. As a result, the process of guessing is partially or completely replaced by an adaptation rule (Fig. 2). The conducted literature research has revealed a great variety of adaptation methods that allow to estimate the covariance matrices \( \hat{Q}_k \) and \( \hat{R}_k \). Among them is first the maximum likelihood estimator (MLE) [14, 15] which has been adopted by several authors in the past, e.g. MEHRA [14, 16], MAYBECK [15], MOHAMED and SCHWARZ [17], VEPA [18], OUSSALAH [19], and SONG [20]. Furthermore, the MIT-rule [20] and the master-slave approach [21, 22], the
multiple model estimation [23], covariance matching [24] and correlation techniques are known as noise adaptation principles. A general overview is provided in [25].

Since there is usually more knowledge available about the measurement noise, we focus in this paper on estimation of process noise which is often much more opaque and intransparent for the practitioner. Due to the results of a preliminary study and for sake of brevity, two promising approaches have been selected as representatives for adaptive Kalman filters: the master-slave KF (MS-CKF) and the Maybeck KF (MB-CKF). Both are either based on processing the innovation $\nu_k$ and its covariance $P_{kk}^{yy}$ or the residual $\hat{x}_k$ and its covariance $\tilde{P}_k$, defined by

$$\hat{x}_k = \hat{x}_k^+ - \hat{x}_k^- = \mathcal{K}_k \nu_k$$

$$\tilde{P}_k = P_k^- - P_k^+ = \mathcal{K}_k P_{kk}^{yy} \mathcal{K}_k^T.$$  

Both adaptation rules use system knowledge provided by the master filter in order to improve on the overall estimator performance. This information is used to compute empirical covariance matrices over defined time windows as it is shown in the following section.

2.3. Process noise estimation using MS-CKF

The MS-CKF is the straightforward concept to estimate process and/or measurement noise covariances when using a SPKF as master filter. The idea is to expand the existing filter by another (linear or nonlinear) KF to estimate directly the master filter’s design parameters. Since there is often no detailed knowledge about the slave filter’s internal model available, a linear KF is the obvious first choice [25]. Then, the following equations hold for the linear slave KF:

$$\hat{x}_{s,k} = \hat{x}_{s,k-1} + Q_{s,d}$$

$$P_{s,k} = P_{s,k-1} + Q_{s,d}$$

$$\hat{y}_{s,k} = C_{s,k} \hat{x}_{s,k} + d_{s,k}$$

$$P_{xy}_{s,k} = P_{s,k} C_{s,k}^T$$

$$P_{yy}_{s,k} = C_{s,k} P_{s,k} C_{s,k}^T + R_{s,d}$$

$$\hat{x}_{s,k} = \hat{x}_{s,k}^- + \mathcal{K}_{s,k} \nu_{s,k}$$

$$P_{s,k}^+ = P_{s,k}^- - \mathcal{K}_{s,k} P_{s,k}^{yy} \mathcal{K}_{s,k}^T$$

$$\nu_{s,k} = y_{s,k} - \hat{y}_{s,k}$$

$$\mathcal{K}_{s,k} = P_{s,k} x_{s,k} \left( P_{s,k}^{yy} \right)^{-1}$$

The index $S$ labels all variables related to the slave filter. The state $x_{s,k}$ contains the critical and unknown process noise parameter. For estimation of process noise $\hat{Q}_k$ based on innovation
\( \mathbf{v}_k \), the above Eqs. (11) must be extended by

\[
\mathbf{d}_{s,k} = \text{diag}\left\{ \mathbf{K}_k^T \left( \mathbf{P}_k^+ - \hat{\mathbf{Q}}_{k-1} - \mathbf{P}_k^+ \right) (\mathbf{K}_k^T)^T \right\}, \quad \mathbf{C}_{s,k} = \mathbf{K}_k^T \circ \mathbf{K}_k^T \tag{12}
\]

\[
\mathbf{y}_{s,k} = \text{diag}\left\{ \frac{1}{N} \sum_{i=k-N+1}^k \mathbf{v}_i \mathbf{v}_i^T \right\}, \quad \hat{\mathbf{Q}}_k = \text{diag}\left\{ \hat{\mathbf{x}}_{s,k}^+ \right\} \quad \tag{13}
\]

wherein \( \mathbf{K}_k^T \) is the Moore-Penrose inverse of the Kalman Gain defined by \( \mathbf{K}_k^T = (\mathbf{K}_k^T \mathbf{K}_k)^{-1} \mathbf{K}_k^T \). The operator \( \text{diag}\{\cdot\} \) is employed for generating a diagonal matrix from a given column vector as well as for extracting the diagonal elements of a square matrix as a column vector. The symbol \( \circ \) indicates an element-wise multiplication. For estimation of process noise based on state residual \( \hat{\mathbf{x}}_k \), the Eqs. (12) - (13) must be replaced by

\[
\mathbf{d}_{s,k} = \text{diag}\left\{ \mathbf{P}_k^+- \hat{\mathbf{Q}}_{k-1} - \mathbf{P}_k^+ \right\}, \quad \mathbf{C}_{s,k} = \mathbf{I} \quad \tag{14}
\]

\[
\mathbf{y}_{s,k} = \text{diag}\left\{ \frac{1}{N} \sum_{i=k-N+1}^k \hat{\mathbf{x}}_i \hat{\mathbf{x}}_i^T \right\}, \quad \hat{\mathbf{Q}}_k = \text{diag}\left\{ \hat{\mathbf{x}}_{s,k}^+ \right\} \quad \tag{15}
\]

One advantage of the residual-based approach is that no computation of the pseudo inverse \( \mathbf{K}_k^T \) is necessary [26]. The measurement information \( \mathbf{y}_{s,k} \) for the slave KF stems from either the diagonal elements of the empirical innovation covariance matrix \( \mathbf{P}_k^{\text{by}} \) or the diagonal elements of the empirical residual covariance matrix \( \hat{\mathbf{P}}_k \). The choice of an adequate window length \( N \) depends on how fast it is intended to compensate and adapt wrong noise assumptions. Furthermore, the initial values of the slave filter \( \hat{\mathbf{x}}_{s,0}^+ \) and \( \mathbf{P}_{s,0}^+ \) as well as the (constant) slave covariance matrices \( Q_{s,k} = Q_{s,d} \) and \( R_{s,k} = R_{s,d} \) must be chosen carefully (see Sec. 4). For practical usability, both process and measurement noise have been implemented in a joined framework.

3. Dynamic modeling and observability assessment

This section introduces the filter models obtained from physical white-box modeling. A simplified rotor-nacelle model is introduced which is employed for estimation of nacelle dynamics, wind speed and tower eigenfrequency. Moreover, a brief discussion on observability of system states and parameters using empirical Gramian matrices is included.

3.1. Simplified dynamic wind turbine model

For simple estimation and control tasks it is advantageous to have basic design models at hand. These are suited as low-order internal models and provide insight into the system dynamics with only two nonlinear differential equations. The first dynamic equation comes from the conservation of angular momentum. Applied to the wind turbine’s drive-train, it yields

\[
\dot{\varphi}_g = \frac{\varrho}{2} \frac{\pi R^3}{\Theta_x + \Theta_{sg}^2 \varrho_{gb}} C_M(\lambda, \beta) \left( v_w - \dot{x}_T \right)^2 - \frac{i_{gb}}{\Theta_x + \Theta_{sg}^2 \varrho_{gb}} M_g \quad \text{with} \quad \lambda = \frac{\varrho g R}{v_w - \dot{x}_T} \tag{16a}
\]

where \( \lambda \) is the so called tip-speed-ratio. The second dynamic equation is obtained from the conservation of linear momentum for the nacelle in axial direction. It reads

\[
\ddot{x}_T = \frac{\varrho}{2} \frac{\pi R^2}{m_T} C_T(\lambda, \beta) \left( v_w - \dot{x}_T \right)^2 - 2 \zeta_{tx} \omega_0 \dot{x}_T - \omega_0^2 x_T \quad . \tag{16b}
\]

The model has two control inputs, the generator torque \( M_g \) and the collective pitch angle \( \beta \) as well as one disturbance input, the effective wind speed \( v_w \). These Eqs. (16) are valid in the partial and in the full load regime which comes along with complex two-dimensional characteristics for coefficients of torque \( C_M \) and thrust \( C_T \), though.

4
This section describes the static and noise adaptive design of the employed SPKF for the proposed problem solution.

3.2. Observability assessment

The concept of observability is useful to assess certain properties of dynamic state-space systems in a systematic fashion. By conducting an observability and/or identifiability analysis the control engineer attempts to evaluate the amount of information about a particular quantity in the measurable output $y$. Considering linear systems, observability is evaluated for instance with Kalman’s observability criterion or the observability Gramian matrix [27]. However, these well-known approaches cannot be applied to nonlinear systems directly. Thus among others, empirical observability Gramian (EOG) matrices have been introduced in [28] which provides a powerful tool to perform such an analysis based on either measurement or simulation data. The empirical observability Gramian is a nonlinear analogue to the linear observability Gramian defined in [29], and has been proposed as a method for measuring the local observability of nonlinear systems in [30]. For sake of brevity, we can only provide a quick glimpse into EOG analysis here since a detailed observability study is out of the paper’s scope.

As illustrative example we have computed the EOG for the model (16) where the wind speed $v_w$ and the eigenfrequency $\omega_0$ have been augmented as constant states, yielding the state vector

$$\mathbf{x} = \begin{bmatrix} \dot{\phi}_g & x_T & x_T & v_w & \omega_0 \end{bmatrix}^T.$$  \hfill (17)

The EOG has been assessed for time windows of $T_0 = 1$ s with $N = 10$ samples each and its inverse condition number (ICN) is evaluated as indicator for observability which is shown in Fig. 3 (left plot). A low ICN indicates low observability and vice versa. However, this only describes the complete system and cannot be associated with certain states directly. Thus, we used additionally the median singular values (over a simulation time $T = 600$ s) and then related them to the individual states $x_1, x_2, \ldots, x_5$ using the singular vectors of the EOG. The results are displayed in Fig. 3 (right plot). As expected, when measuring the complete state with configuration $y_{(11)}$ the observability is very high. For configuration $y_{(21)}$, where only the generator speed and nacelle acceleration are measured, the observability is still reasonable. The reduction in observability compared to $y_{(11)}$ is mainly influenced by the unknown wind speed $x_4$. The configurations $y_{(31)}, y_{(41)}$ and $y_{(51)}$ are rated as non-observable since one or more states have an observability index $g(x_i) \leq 10^{-6}$ which has been found as lower bound from simulation studies with various filter configurations.

4. Static and adaptive filter design

This section describes the static and noise adaptive design of the employed SPKF for the proposed problem solution.
4.1. Optimal static filter design

The most efficient way to find the best static filter parameters is to establish a mathematical optimization problem. The cost function for this design optimization is defined as

$$ J(z) = J_x(z) + J_y(z) $$

(18)

where $z$ represents the optimization variables. The first part of the cost function $J_x$ penalizes the deviations from the true state vector $x$ and the second part penalizes the deviations from the measured output $y$. We propose the following cost functions for the Kalman filter:

$$ J_x(z) = \sum_{i=10}^{N} \left( x_i - \hat{x}_i(z) \right)^T \left[ \begin{array}{cc} W_x & 0_n \\ 0_n & W_{\hat{x}} \end{array} \right] \left( x_i - \hat{x}_i(z) \right) $$

(19)

$$ J_y(z) = \sum_{i=10}^{N} \left( y_i - \hat{y}_i(z) \right)^T \left[ \begin{array}{cc} W_y & 0_q \\ 0_q & W_{\hat{y}} \end{array} \right] \left( y_i - \hat{y}_i(z) \right) $$

(20)

using the optimization variables

$$ z = [ q_d^T \ r_d^T]^T $$

(21)

where $q_d = \text{diag}\{Q_d\}$ and $r_d = \text{diag}\{R_d\}$ are the diagonal elements of the covariance matrices. The matrices $W_x$, $W_{\hat{x}}$, $W_y$ and $W_{\hat{y}}$ are positive semi-definite weighting matrices that are predefined in order to put emphasis on the states, outputs and/or its time-derivatives. Eventually, the optimization problem (OP) for the KF is stated as follows:

$$ \min_z J(z) = J_x(z) + J_y(z) $$

s.t. $[\hat{x}_i^+, \hat{y}_i] = \text{CKF}(z) : \forall i$

$$ \bar{z} \leq z \leq \overline{z} $$

(22)

Therein, $z$ are the lower and $\overline{z}$ the upper bound constraints. The optimization is carried out for S representative scenarios with $N$ time samples each. A scenario uses data generated either directly from FAST8 [31] simulation or from design model simulation. This data set includes time trajectories of $x$, $d$, $u$ and $y$. Only the last two quantities are (obviously) used by the SPKF to compute the estimates $\hat{x}_i^+$ and $\hat{y}_i$ in each iteration.

The nonlinear OP with constraints of Eqs. (22) is set up in Matlab and solved by an interior point algorithm [32]. The lower and upper bounds for the optimization variables as well as step and optimality tolerances are selected so that a reasonable trade-off between computation time and accuracy is achieved. In order to determine feasible initial values for the optimization variables, a coarse pre-optimization based on two approaches is considered. The first approach consists of generating and simulating various sets of scenario configurations and selecting scenarios that present the best performance regarding the cost function $J(z)$. The second approach generalizes the aforementioned idea by implementing a genetic algorithm, using $J(z)$ as fitness function.

An illustrative scenario for this parameter optimization is shown in Fig. 4 where the weighting matrices were set to $W_x = W_{\hat{x}} = I_n$ and $W_y = W_{\hat{y}} = 0_q$ for sake of simplicity. This optimization approach provides the optimal configuration for the standard SPKF which also defines the initial values of the adaptive Maybeck and master-slave filter.

4.2. Noise adaptive design approach

The noise adaptation is supposed to tackle changing and/or initially unknown noise statistics. These are determined not only by white Gaussian input and output noise but also by changing
model errors and model uncertainties like changing average wind speed and turbulence. The design procedure is similar to the static design. Once an optimal filter with static parameters is found, the optimization problem (22) is formulated for optimizing the adaptation-rule-specific set of optimization variables. For the MS-CKF, the optimization variables are given by

\[
\mathbf{z} = \begin{bmatrix} \mathbf{q}_s^T \\ \mathbf{r}_s^T \end{bmatrix}^T
\]

where \( \mathbf{q}_{s,d} = \text{diag}\{\mathbf{Q}_{s,d}\} \) and \( \mathbf{r}_{s,d} = \text{diag}\{\mathbf{R}_{s,d}\} \). Fig. 5 shows exemplarily the adaptive wind speed estimation with and without optimized slave parameters. Despite a strong sensor noise environment, the SPKF with optimized slave parameters adapts to the wind trajectory accurately by adjusting the wind speed’s variance. The same applies for the predicted output.

5. Simulation studies and results
Finally, after introducing the necessary algorithms, models and the filter design, the obtained simulation results are presented. First, the chosen scenario configurations are presented in order to generate a preferably realistic set of test data. Then, the considered estimation problems (EP) are defined and obtained simulation results are discussed.
5.1. Configuration of test scenarios and simulator details

The test scenarios are defined to investigate the standard and noise adaptive filter performance under realistic conditions. Realistic conditions mean that

(i) sensor data (both, control inputs and system outputs) is corrupted by white Gaussian noise sequences with low, medium and high noise levels,

(ii) sensor drift, sensor errors and faulty data are excluded from the study,

(iii) wind fields with average wind velocities ranging from 6 to 21 m/s are tested,

(iv) turbulence intensities from 1 to 15% and influence of wind shear are present,

(v) simulation data from FAST8 [31] as well as from different control oriented design models,

(vi) tower and blade dynamics are released while the substructure dynamics have been excluded.

The simulation studies have been conducted with a standard industrial wind turbine controller. Thus, two separate single-input-single-output (SISO) control loops are used to compute the generator torque and the collective pitch angle, both combined controlling the generator speed. The generator torque follows an optimal torque curve while for the collective pitch angle a gain-scheduling PID-controller is applied, similar to [33].

5.2. Defined estimation problems

The complete estimation problem for wind turbine application has been introduced and (almost completely) solved in [5], except for the noise adaptation of the filter parameters. This paper closes the gap and demonstrates the approach for two designated problems. The first problem (EP1) addresses the simplified state and process noise estimation in order to reconstruct the effective wind speed estimation under different noise conditions. The sensor equipment considered for wind speed estimation solely includes

\[ u = \begin{bmatrix} M_g \\ \beta \end{bmatrix}^T \]

and

\[ y = n_g, \]

therefore generator torque and speed as well as collective pitch angle are known to the filter. With this set-up the incoming wind speed, the rotor angular speed and the rotor torque are estimated.

The second problem (EP2) addresses the combined wind speed and nacelle fore-aft estimation, both relevant for wind turbine control. The sensor equipment comprises

\[ u = \begin{bmatrix} M_g \\ \beta \\ \vec{x}_r \\ n_g \end{bmatrix}^T \]

and

\[ y = \begin{bmatrix} \dot{x}_t \\ n_g \end{bmatrix}^T. \]

The following items are part of EP2: The adaptive estimation of effective wind speed without nacelle anemometer, the state estimation of nacelle movement only based on inertial measurement units (IMU), the reconstruction of rotor torque and rotor thrust without load measurements, and the online estimation of process noise covariances for a good filter parametrization. This estimation problem is solved by an adaptive MS-CKF and a MB-CKF using the simplified rotor-nacelle model of Eqs. (16) as filter model.

5.3. Estimation results

First, the EP1 is considered. Exemplary simulation results are shown for the noise-adaptive wind estimation in Fig. 6 where the optimal standard SPKF has been designed for a too low noise environment. However, this design results in an increasingly aggressive filter behaviour when sensor noise increases. Contrary, both adaptive filters can manage the increased noise level and improve their performance over time.

Secondly, the EP2 is investigated where a conservative choice of the process noise covariance

\[ q_{ii} = 10^{-4} \]

with \( i = \{1, 2, 3, 4\} \) is initially made (cf. Fig. 7). This design results in poor estimation quality. The simulation starts at \( t = 0 \) s without adaptation and at \( t = 300 \) s both adaptation rules are suddenly activated to improve on the filter performance. As hoped for, the noise parameters are immediately adapted after switch-on and reduce the state estimation errors. Interestingly, the MS-CKF outperforms the CKF and also the MB-CKF for the states \( x_t \) and \( v_w \). This has mainly two reasons: First, the Maybeck estimator has only the window size
Figure 6. Process noise adaptation comparing CKF, MB-CKF and MS-CKF state estimates with the true states and the process noise estimate.

Figure 7. Process noise adaptation comparing MB-CKF state estimates and MS-CKF with the true states (left) and the process noise estimates (right).

N as design parameter. Consequently all parameters are adapted more or less in the same magnitude order since there is less flexibility. Secondly, the noise parameters $q_{11}$ and $q_{44}$ both affect the innovation $v_k$ and its covariance matrix $P_{yy}$ similarly. Thus, both should not be estimated simultaneously. With the master-slave approach the relevant noise parameters can be
specifically selected for adaptation which is a relevant attribute for practical realization. Hence, the MS-CKF sets the first noise parameter $q_{11}$ constant which results in a significant performance gain especially for the wind estimate.

5.4. Discussion and recommendations
The choice of the free design parameters often remains a mystery of Kalman filtering. Still, this paper demonstrates that there are adaptive estimation techniques as alternative to circumvent the process of guessing the covariance matrices. We propose the following steps:

(i) Conduct a numerical optimization for the static filter parameters considering time-varying sensor noise for control inputs and outputs.
(ii) Carry out an automated testing with a preferably complete set of scenarios to test the design parameters from optimization.
(iii) Perform a parameter optimization for the adaptive filter with the previously found optimal static design.
(iv) Assess the adaptive filter’s performance against the static design and evaluate the limitations and limits of the noise adaptive filter parametrization.

One important outcome of this research study is, not to adapt all noise parameters at the same time. Hence, the practitioner is encouraged to introduce his or her a priori expert knowledge and pool it conveniently with the automated estimation approach. The master-slave technique has proven to be an appealing approach since only the critical parameters are released for estimation. The computational effort increases only moderately compared to the standard SPKF variants.

6. Conclusions and outlook
In this contribution, standard and adaptive nonlinear sigma-point Kalman filters have been investigated for wind turbines in order to tackle the problem of good filter parametrization with changing noise statistics. The relevant algorithms and dynamic wind turbine models have been presented and design suggestions have been made. The master-slave filter has been found to be a suitable algorithm for the adaptation of the SPKF.

In a nutshell, this paper demonstrates how adaptation methods exploit the additional information provided by the SPKF to improve the accuracy of the wind turbine’s state estimates. Data-based adaptation techniques are an essential feature to cope with time-varying noise statistics and poor initial design in practical realization. Future work will focus on a detailed observability and identifiability analysis for model and filter parameters to assess the limits of adaptive Kalman filtering. Moreover, multiple model estimation techniques will be investigated as complement to the master-slave approach.

References
[1] Bossanyi E, Savini B, Iribas M, Hau M, Fischer B, Schlipf D, van Engelen T, Rossetti M and Carcangiu C E 2012 *Wind Energy* 15 119–145
[2] Schlipf D, Schlipf D J and Kühn M 2013 *Wind Energy* 16 1107–1129
[3] Gros S and Schild A 2017 *International Journal of Control* 90 2799–2812
[4] Schild A, Gros S and Ritter B 2016 High performance nonlinear model predictive control for wind turbines *Proceedings of Wind Europe Summit 2016*
[5] Ritter B, Schild A, Feldt M and Konigorski U 2016 *Journal of Physics: Conference Series* 753 052029
[6] Ritter B, Schild A and Konigorski U 2016 Making nonlinear state estimation techniques ready for use in industrial wind turbine control systems *Proceedings of Wind Europe Summit 2016*
[7] Kalman R E 1960 Trans. ASME - Journal of Basic Engineering 82 35–45
[8] Brammer K and Siffling G 1989 *Kalman-Bucy Filters* (Artech House, Inc.)
[9] Simon D 2006 *Optimal State Estimation. Kalman, H∞, and Nonlinear Approaches* (John Wiley & Sons)
[10] van der Merwe R and Wan E A 2001 *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP ’01).* 2001 6 3461–3464
\[ x_k \] dynamic state vector
\[ u_k \] control input vector
\[ d_k \] disturbance input vector
\[ \theta_k \] model parameter vector
\[ y_k \] measured output vector
\[ y_k^p \] predicted output vector
\[ w_k \] process noise vector
\[ n_k \] measurement noise vector
\[ f(\cdot) \] discrete-time process model
\[ h(\cdot) \] measurement model
\[ z \] optimization variables
\[ J(z) \] general cost/objective function
\[ J_0, J_y \] cost function for states/outputs
\[ W, W_0 \] weighting matrices for states
\[ W_y, W_u \] weighting matrices for outputs
\[ \theta_{\text{ub}}, \theta_{\text{lb}} \] unity/zero matrix dim \([i \times j]\)
\[ k \] sample step \( t(k) = KT \)
\[ T_0 \] sample time
\[ n \] number of states
\[ p, q \] number of inputs/outputs
\[ N \] number of samples
\[ S \] number of scenarios
\[ \hat{x}_k \] a priori/posterior state estimate
\[ w_k, v_k \] state residual, meas. innovation
\[ Q_k, R_k \] process noise covariance matrix
\[ R_k, R_{\text{d}} \] measurement noise cov. matrix
\[ y_k, r_k \] diag. elements of \( Q_k, R_k \)
\[ \hat{Q}_k, \hat{R}_k \] estimates of noise cov. matrices
\[ K_k \] Kalman Gain and its pseudo inverse
\[ P_{k-1}^n, P_k^n \] innovation/cross covariance matrix
\[ \hat{P}_k \] residual error covariance matrix
\[ X_{\text{d}}^{-1} \] before passing through \( f(\cdot) \)
\[ X_{\text{d}} \] after passing through \( f(\cdot) \)
\[ X_{\text{p}} \] before passing through \( h(\cdot) \)
\[ X_{\text{a}} \] after passing through \( h(\cdot) \)
\[ Y_k, Y_{\text{d}} \] matrix w/ \( 2n \) col. vectors of \( x_k \)
\[ Z_k, Z_{\text{d}} \] matrix w/ \( 2n \) col. vectors of \( x_k \)
\[ A_k, A_{\text{d}} \] a priori/posterior state est. (slave filter)
\[ P_{\text{sl}} \] equivalent tower top mass
\[ \Theta_0 \] rotor inertia (high-speed-side)
\[ \Theta_\theta \] generator inertia (low-speed-side)
\[ \lambda \] blade-tip speed ratio
\[ \beta \] collective blade pitch angle
\[ v_w \] hub-height wind speed
\[ \omega_0 \] first tower eigenfrequency
\[ \zeta \] modal damping ratio
\[ \varrho \] air mass density
\[ R \] blade tip radius
\[ C_T \] aerodynamic thrust coefficient
\[ C_M \] aerodynamic torque coefficient
\[ \lambda \] blade-tip speed ratio
\[ \beta \] collective blade pitch angle
\[ \omega_0 \] first tower eigenfrequency
\[ \zeta \] modal damping ratio
\[ \varrho \] air mass density
\[ R \] blade tip radius
\[ C_T \] aerodynamic thrust coefficient
\[ C_M \] aerodynamic torque coefficient
\[ \lambda \] blade-tip speed ratio
\[ \beta \] collective blade pitch angle
\[ v_w \] hub-height wind speed
\[ \omega_0 \] first tower eigenfrequency
\[ \zeta \] modal damping ratio
\[ \varrho \] air mass density
\[ R \] blade tip radius
\[ C_T \] aerodynamic thrust coefficient
\[ C_M \] aerodynamic torque coefficient
\[ \lambda \] blade-tip speed ratio
\[ \beta \] collective blade pitch angle
\[ v_w \] hub-height wind speed
\[ \omega_0 \] first tower eigenfrequency
\[ \zeta \] modal damping ratio

**Table 1. Nomenclature.**