COMPARISON OF DIFFERENT KALMAN FILTERS FOR TRACKING NONLINEAR TRANSMISSION TORQUES

Geir Hovland *

* *School of Information Technology and Electrical Engineering,*
The University of Queensland
Brisbane QLD 4072
Australia

Abstract: This paper presents parameter estimation of physical time-varying parameters for a mechanical system with a nonlinear gear-box transmission torque. Four different variants of the Kalman filter are compared extensively on models of different complexities. The first filter is the well-known Extended Kalman filter. The second and third filters are recent (2000 and 2001) developments of unscented versions of the Kalman filter. The fourth filter is a new (2000) and compact version of the adaptive sensitivity-based Kalman filter. The four different approaches have different complexities, behaviour and advantages that are surveyed in this paper. The paper shows that some of the filters are able to successfully track a rapidly changing transmission torque.

Keywords: Identification, transmission torque, hysteresis

1. INTRODUCTION

Identification is a field of systems and control that has produced powerful methods and tools for modelling based on collected data. While the current status of the linear identification theory is a mature topic, challenges exist for identification of more complex systems (nonlinear, distributed, hybrid, large scale physical models), in particular goal-oriented modelling issues (identification for control, diagnosis, detection, monitoring, maintenance optimisation, etc).

In this paper a survey and evaluation of four different methods for parameter estimation of a mechanical system with a nonlinear gear-box transmission torque is presented. The main goal of testing the different methods for parameter tracking, is to find a reliable tool for diagnostics of gear-box transmissions. When identification is goal-oriented, certain requirements are usually placed on the algorithms. First, they must be able to identify parameters in physical models. Second, due to wear and tear or mechanical failure, the algorithms must be able to track relatively fast changes in the physical parameters. The four parameter estimation filters discussed in this paper are all variants of the Kalman filter that meet these requirements.

The use of Kalman filtering for estimating nonlinear transmission torques is not new. Examples were presented by (Hovland *et al.*, 2002; Lagerberg and Egardt, 2003). However, the author believes that estimation of a nonlinear transmission torque with hysteresis presented in this paper is new. The author has previously presented experimental results for estimating the transmission torque for an industrial robot based on the extended Kalman filter. In this paper, however, the focus is on evaluating and comparing different methods for nonlinear estimation. Hence, knowledge of the 'true' transmission torque is needed to evaluate the parameter estimation errors of the different filters. In this paper, the four versions of the Kalman filter were tested on simulated transmission models both with and without hysteresis in presence of both measurement and process noise. The approach of using of a 'true' simulated system with added noise for evaluating the effectiveness of the Kalman filter for estimating backlash size was also chosen by (Lagerberg and Egardt, 2003).

2. MODEL

Figure 1 illustrates backlash in a gear-box transmission. Depending on the stiffness of the mating gears, the size of the contact areas, lubrication, etc., the transmitted torque as a function of the positional difference between the gears can take on a variety of different forms. In this paper the actual functional dependency of the torque vs. gear positions is assumed completely unknown to the different Kalman filter techniques.

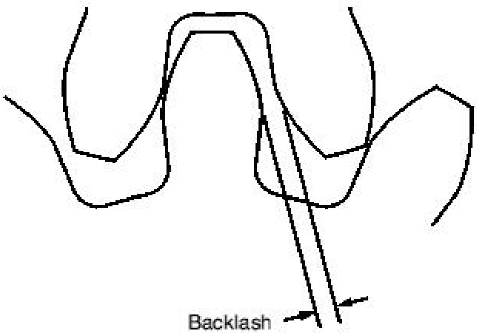

Fig. 1. *Illustration of backlash in a gear-box transmission.*

The differential equations of motion for a two-mass motor-arm mechanical system with an unknown transmission torque τ_K are given as follows.

$$\begin{aligned}
\dot{x_1} &= x_3 \\
\dot{x_2} &= x_4 \\
\dot{x_3} &= \frac{1}{m_1}\left(u - D(x_3 - x_4) - \tau_K - \tau_f - K_r\right) \\
\dot{x_4} &= \frac{1}{m_2}\left(D(x_3 - x_4) + \tau_K\right) \\
\tau_f &= C_v x_3 + C_c \frac{2}{\pi} \operatorname{atan}(\alpha x_3)
\end{aligned} \tag{1}$$

where m_1, m_2 are the motor and arm inertias, respectively. τ_K is a nonlinear backlash and elasticity function to be identified. K_r is a P-controller on the motor position, u is an additional torque input used for identification, D is a known damper coefficient and τ_f is a motor friction torque. The model is written on short form as follows

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{p}) + v$$

where $p = [\tau_K]^T$ is the augmented state and the unknown nonlinear transmission torque to be estimated. τ_f is the viscous and Coloumb friction torque model. To get the Kalman filter to work properly, a continuous approximation (here the atan function) of the Coloumb friction is needed in order to compute the gradients $\frac{df}{dx}$ at speeds close to zero.

The measurements are motor and arm positions, ie. the measurement function $\mathbf{h}()$ equals

$$y = \mathbf{h}(\mathbf{x}, \mathbf{u}, \mathbf{p}) + w = [x_1, x_2]^T + w$$

where v, w are noise processes. The model is illustrated in Figure 2.

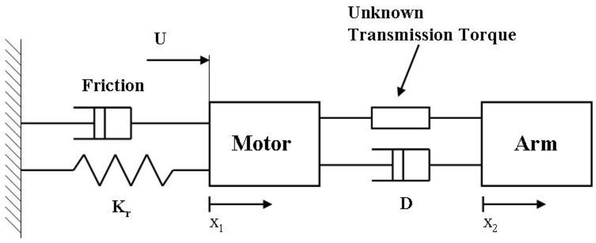

Fig. 2. *Mechanical system with unknown transmission torque.*

One advantage of the state-augmented filter approach, is the fact that no particular structure of the transmission torque has to be assumed. Instead, the torque τ_K is added to the differential equations as an augmented state. This feature is extremely useful, since the same model can be used for a large variety of different systems; for estimation based on a 'true' simulated system to estimation on measured data where the true structure may be completely unknown. The concept of augmented states is explained in more detail in the following section.

3. AUGMENTED STATES

The augmented states are introduced into the system equations in the following way.

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{p}) \\ 0 \end{bmatrix} + v \tag{2}$$

where $\mathbf{p}$ is an unknown parameter vector. The differential equation for $\mathbf{p}$ is a random-walk process due to the process noise v. The Kalman filter equations will treat $\mathbf{p}$ as a parameter vector that can change value at every parameter update step. Hence, depending on the tuning of the Kalman filter, the parameter $\mathbf{p}$ can model both slow and fast changing dynamics. It is important to note that no structure of the underlying dynamics has to be assumed. For the unknown transmission torque problem treated in this paper, $\mathbf{p}$ is simply chosen to be the transmission torque τ_K. The transmission torque can contain both nonlinearities and memory effects such as hysteresis. The state-augmented Kalman filter will at every time step calculate the transmission torque that best matches the measurements and the model. After the identification procedure has been completed, the state-augmented transmission torque is plotted on the y-axis versus the position difference $x_1 - x_2$ on the x-axis. If the system has been properly exited, then this plot will reveal the structure of the transmission torque. If necessary, a structured model of the transmission torque can then finally be estimated from this plot.

The complete augmented model can then be written as follows.

$$
\begin{aligned}
\dot{x_1} &= x_3 + v_1 \\
\dot{x_2} &= x_4 + v_2 \\
\dot{x_3} &= \frac{1}{m_1}\left(u - D(x_3 - x_4) - p - \tau_f - K_r\right) + v_3 \\
\dot{x_4} &= \frac{1}{m_2}\left(D(x_3 - x_4) + p\right) + v_4 \\
\dot{p} &= v_5
\end{aligned}
$$

where v_i variables are noise processes acting on the state and parameter variables. The covariance matrix for the augmented noise process v is given as

$$
E(v(t)v^T(t^{'})) = \begin{bmatrix} \mathbf{Q}_{xx} & 0 \\ 0 & \mathbf{Q}_{pp} \end{bmatrix} \delta(t - t^{'}) \quad (3)
$$

where $\mathbf{Q}_{xx}$ is the covariance of the noise on the state derivatives, while $\mathbf{Q}_{pp}$ is the covariance of the noise on the augmented state derivative, ie. the derivative of the backlash and spring force. $\mathbf{Q}_{pp}$ can be seen as the main tuning parameter of the augmented filter. If $\mathbf{Q}_{pp}$ is set to zero, the parameter $\mathbf{p}$ will remain a constant equal to its initial value. The larger the value of $\mathbf{Q}_{pp}$, the faster the augmented state will be updated.

4. NONLINEAR PARAMETER TRACKING

This section gives a brief description of the different versions of the Kalman filters that are surveyed in this paper and where to find descriptions of the implementation details. All the filters were implemented using the Newmat C++ Matrix library, (Davies, 2004).

4.1 *State-Augmented Extended Kalman Filter*

The state-augmented extended Kalman filter (EKF) is the extension of the standard Kalman filter to nonlinear systems. At each time step the dynamics are linearised at the current estimates of $\mathbf{x}$ and $\mathbf{p}$. A compact notation of the EKF that is easily implemented in Newmat can be found in (Bohn, 2000). The advantages of the EKF is the ease of implementation and the relatively good filter performance at low computational costs. However, the EKF has two main drawbacks: 1) the linearisation of the nonlinear dynamics is only an approximation and the filter can lead to biased parameter and state estimates for strongly nonlinear systems. 2) The EKF requires the first order gradients of the models, ie. the partial derivatives of the system equations $\mathbf{f}()$ and $\mathbf{h}()$ with respect to the state vector $\mathbf{x}$ and the parameter vector $\mathbf{p}$. These derivatives can either be found numerically or calculated analytically for the specific model. In the work presented here, all partial derivatives (including the higher order derivatives for the AEKF) were computed numerically.

4.2 *Unscented Kalman Filter*

The Unscented Kalman Filter (UKF) was developed by (Julier and Uhlmann, 1997; Julier *et al.*, 2000). The UKF's approach to generalise the nonlinear system is different from that chosen in the EKF. Instead of approximating a nonlinear function by first order gradients, the UKF approximates probability distributions by the use of the Unscented Transform. While the EKF calculates the Jacobian and solves a Ricatti equation to estimate the covariances, the UKF uses so-called sigma points and propagates all of them through the system model to numerically estimate the state vector $\mathbf{x}$ and the output vector $\mathbf{y}$ as well as their covariances. Note, there are different versions of the UKF. As an alternative, the process (and even the measurement) noise could also be included in the transformed vector (Wan and van der Merwe, 2000; Wan and van der Merwe, 2001*a*). This is recommended in the case where the process noise is state-dependent.

4.3 *Square-Root Unscented Kalman Filter*

A major drawback of the UKF is its vulnerability for numerical instability. In particular, the positive definiteness of the state estimation error covariance matrix can be lost in the update step of the UKF. To overcome this flaw, the Square-Root Unscented Kalman Filter (SRUKF) was proposed by (Wan and van der Merwe, 2001*b*). Similar to the standard square-root Kalman filter, the Cholesky factor of the state estimation error covariance matrix, $\mathbf{S}^a$, is calculated and updated, which can be directly used to determine the sigma points for the next propagation step. Therefore, the determination of the sigma points is far less complex because there is no need for a Cholesky factorisation. However, the algorithm contains other operations of cubic complexity (e.g. the QR decomposition) such that the complexity of the SRUKF and the UKF are both of the same order ($O(L^3)$). Due to the improved numerical properties the SRUKF is more robust than the UKF.

4.4 *Adaptive Extended Kalman Filter*

It is shown in (Ljung and Söderström, 1993; Ljung, 1999) that the EKF can be rewritten and interpreted as a recursive prediction error method with some simplifications in the computation of the gradient. The simplifications do not only lead to poor convergence properties, they also make it impossible to include elements of the noise covariance matrices ($\mathbf{Q}$ and $\mathbf{R}$) in the set of unknown parameters, (Ljung, 1979). The Adaptive Extended Kalman Filter (AEKF) alleviates these problems at the cost of computational burden.

A new compact matrix notation for the AEKF was developed by (Bohn, 2000). Instead of element-wise derivatives Bohn introduced a notation using Kronecker products, matrix stacking operations and matrix derivatives. Hence, implementation of the AEKF becomes straightforward for models independent of size by using a modern matrix library such as Newmat.

The main source of complexity of the AEKF algorithm is the propagation of the so-called sensitivities (second order partial derivatives), which grow quickly with the number of states and parameters.

5. TRUST INDICATORS

When it comes to applications of parameter estimation algorithms a crucial issue is to be able to assess the quality and/or reliability of the estimates. As often in control theory, while for linear systems there are a number of tests which can be used, for general nonlinear systems this is a largely unsolved problem. However, despite the lack of a consolidated useful mathematical theory, a number of semi empiric trust indicators can be developed to address these problems. Clearly, for this same reason the usefulness of these indicators in each concrete situation must be tested extensively and must be decided on a case by case basis.

In the sequel we will present three trust indicators we think can be useful in the applications we face. Clearly this list is not unique as it can be extended as concrete application knowledge flows into the indicator design. Also note that the indicators are rather generic, which means that they do not depend on the type of filter in use but rather check certain system-specific observability properties.

5.1 *Observability indicators*

The leading idea is to check whether the (nonlinear!) plant model equations allow for infinitely many solutions at a given operating point. Assume that the parameters are estimated at a given equilibrium point. Then the following equations must be satisfied.

$$\mathbf{0} = \mathbf{f}(\mathbf{x}, \mathbf{p}) \tag{4}$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}, \mathbf{p}) \tag{5}$$

where the $N_x + N_p$ dimensional vector function $\mathbf{f}$ models the plant dynamics and the N_y dimensional vector function $\mathbf{h}$ denotes the available measurements. One important question to answer here is the following. Are there enough measurements to asses $(\mathbf{x}, \mathbf{p})$ from knowledge of $\mathbf{y}$? Mathematically speaking not enough measurements means that infinitely many combinations $(\mathbf{x}, \mathbf{p})$ satisfy the equilibrium equations.

A sufficient condition which excludes this from happening is the following: the Jacobian of the $N_x+N_p+N_y$ vector function $[\mathbf{f}^T\mathbf{h}^T]^T$ must be nonzero. Thus, the following condition is proposed

$$\gamma^{-1}\left(\frac{d}{d[\mathbf{f}^T\mathbf{h}^T]^T}[\mathbf{f}^T\mathbf{h}^T]^T\Big|_{(\mathbf{x},\mathbf{p})=(\hat{\mathbf{x}},\hat{\mathbf{p}})}\right) > \epsilon \tag{6}$$

for a suitable $\epsilon > 0$. $\gamma(\mathbf{A})$ is the condition number of a matrix $\mathbf{A}$ defined as the ratio between the maximum and minimum singular values. See for example (Skogestad and Postlethwaite, 1996) page 87 for a discussion of the use of the condition number in multivariable control.

Several versions of this indicator can be used. For instance, in certain situations it can be useful to check the Jacobian with respect to $\mathbf{p}$ only.

Note that this condition may fail at a number of estimates without consequences on the overall performance. However, frequent failures in this condition indicate potential problems.

5.2 *Prediction error indicators*

The prediction error equals

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}},$$

in words the difference between the real measurement $\mathbf{y}$ and the ones forecasted by the models at current estimates values, is also a source of valuable information for assessing filter current performance.

In particular, it is well known that a filter is performing well if this error is white noise, i.e., $\mathbf{e}$ must have zero mean and no temporal correlation. Checking these conditions is relatively easy and should be done quite frequently in order to have early warning of bad filter performance.

Further, another indicator of filter performance is the covariance of $\mathbf{e}$, which corresponds to the matrix $\mathbf{A}$ of the Kalman filter. Since $\mathbf{e}$ is measurable, $\mathbf{A}$ can be measured as well and its size used to indicate irregularities.

The importance of the prediction error characteristics is stressed by the fact that they can be used for effective tuning of filter parameters like the covariance matrices $\mathbf{Q}$ and $\mathbf{R}$.

As the observability indicator, these magnitudes are generic, i.e. they do not depend on the filter form.

5.3 *Covariance indicator*

Computing covariances of the estimates is also relevant in this context. Indeed, growing covariances indicate that the estimates variance is becoming unacceptably large and thus concrete values are not useful. On the other hand, too small covariance values will prevent estimate updates to take place when relevant new measurements appear. It follows that checking appropriate lower and upper bounds for these magnitudes is crucial for assessing reliability. To a certain extent this indicator is filter dependent.

6. CASE STUDIES

In this section two different case studies are presented. First, estimation of a transmission torque with backlash and a hysteresis effect using the four different filters is presented. The filters are compared based on computational efforts and prediction errors. The prediction errors are measured both on the measurements and on the transmission torque compared to the 'true' system. Second, the same comparisons are made on a nonlinear transmission torque without the hysteresis.

The static models parameters were chosen as follows: $m_1 = 0.006442$, $m_2 = 0.023454$, $D = 0.01$, $C_v = 0.1$, $C_c = 0.27$, $\alpha = 100$, $K_r = 0.01$.

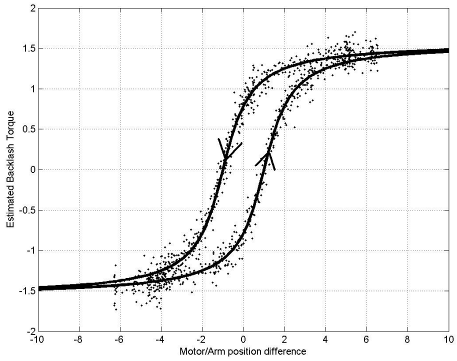

Fig. 3. *Nonlinear transmission torque with hysteresis (solid lines) and function estimate using the state-augmented extended Kalman filter.*

	EKF	UKF	SRUKF	AEKF
CPU	1.00	2.21	2.25	3.44
Pred.Err. e_1	0.064	0.118	0.061	0.159
Pred.Err. e_2	0.086	0.108	0.082	0.123
Parameter Error	1.120	1.138	1.117	1.133

Table 1. *Comparison of filter performance.*

Figure 3 shows the simulated true transmission torque function with solid lines. The memory effect in the transmission torque is illustrated by the arrows in the figure. When the position difference $x_1 - x_2$ is increasing from a low value, the transmission torque follows the lower solid line. If the system reverses direction somewhere on the lower line, the torque will move over to the upper solid line and start decreasing from there. The region between the solid lines is the backlash region. The memory effect depends on the history of the position difference $x_1 - x_2$.

The dots in Figure 3 show the estimated transmission torque when using the state-augmented Kalman filter. A two percent normal distributed noise was added both to the position measurements and to the state estimates when the 'true' system was simulated. The figure clearly shows that the structure of the transmission torque is revealed by the EKF as long as the system is exited in all regions of the position difference $x_1 - x_2$.

Table 1 shows the computational efforts required by the four different filters and the root-mean-square (RMS) prediction errors of: 1) e_1: The motor position x_1, 2) e_2: The arm position x_2 and 3) the estimated transmission torque compared to the true torque. The results show that the EKF is more than twice as fast as any other filter. The most accurate filter is the SRUKF, however the improvements compared to the EKF are only minor. The AEKF was not able to produce any better results than the other filters despite the extra efforts in modelling higher-order effects. One reason for this lack of performance, could be the fact that the transmission torque is changing very rapidly and the adaptive filter is not able to track both the covariance of the unknown parameter and the parameter itself at the same time. For the other filters, the covariance of the unknown parameter $\mathbf{p}$ is kept constant at a higher level than the adapted AEKF covariance. A large covariance means a better ability to track fast changing parameters.

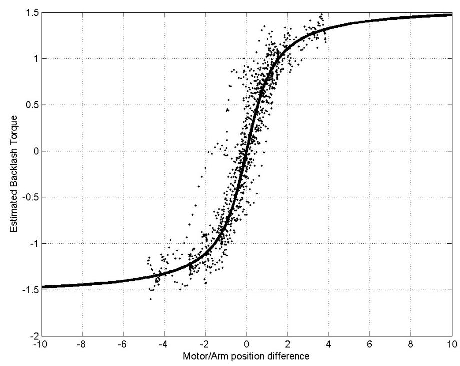

Fig. 4. *Nonlinear transmission torque without hysteresis (solid lines) and function estimate using the state-augmented extended Kalman filter.*

	EKF	UKF	SRUKF	AEKF
CPU	1.00	2.18	2.25	3.46
Pred.Err. e_1	0.059	0.112	0.058	0.098
Pred.Err. e_2	0.066	0.084	0.0063	0.078
Parameter Error	0.832	0.827	0.829	0.828

Table 2. *Comparison of filter performance.*

Figure 4 shows the true simulated transmission torque without the hysteresis effect as a solid line and the estimated torque using the EKF as dots. Again, the filter is successfully able to reveal the unknown structure of the underlying transmission torque with no modelling assumptions made.

Table 2 summarises the results and they are very similar to the results from the model with hysteresis. Again, the EKF is more than twice as fast as the other filters. Again, the SRUKF delivers the best results, however, only marginally better than the EKF. For this model, the AEKF performs better than for the model with hysteresis. This observation also supports the statement above regarding the AEKF's problems with tracking fast changing dynamics. The transmission torque without hysteresis changes less rapidly compared to the torque with the hysteresis.

It should be noted that the tuning of the different filters is a major task. The author has optimised each filter parameter setting individually. The main tuning parameters, the state and measurement covariance matrices, can be quite different from filter to filter. Each filter was tuned to its best performance, and the author believes that the results presented in the Tables 1 and 2 are subjective and a reasonably accurate performance measure of the four filters based on the nonlinear transmission models.

7. CONCLUSIONS

In this paper four different methods for tracking time-varying parameters in nonlinear ordinary differential equations models were presented and compared with respect to computational burden and estimation accuracy. The methods were tested on a mechanical system with a nonlinear transmission torque. Tests were made both with and without hysteresis in the transmission torque.

The extended Kalman filter was the easiest filter to use, both in terms of implementation and in terms of tuning the covariance matrices. For the two case studies presented in this paper, the EKF was also faster than the other approaches. The unscented Kalman filters (UKF and SRUKF) can be seen as a trade-off between the EKF and the more complicated adaptive filter (AEKF). The computational burden is higher than for the EKF, but significantly lower than for the AEKF and the unscented filters produce low prediction errors. The square-root version of the unscented filter produced the highest quality estimates.

The adaptive sensitivity-based version of the Kalman filter was made accessible by (Bohn, 2000). Bohn presented a new compact notation for the AEKF which makes implementation straightforward with a modern matrix library. One complicating factor is the need to compute second-order derivatives. The AEKF was not able to produce results comparable in quality to the other filters for the transmission torque function with hysteresis. For the hysteresis-free model, however, the AEKF produces acceptable results.

The EKF and the SRUKF stand out as the best filters for tracking a rapidly changing nonlinear transmission torque with hysteresis. The SRUKF produces slightly better results than the EKF at the price of a higher computational burden. The author believes that both methods are ideal candidates for monitoring gear-box transmissions during operation. By running the filters regularly over long time periods, the methods provide very useful tools for monitoring degradation of gear-boxes. The tools can also be used to save a significant amount of manual inspection times for manufacturers of gear-boxes when these have to be delivered with a minimum tolerance on backlash and/or stiffness.

8. REFERENCES

Bohn, C. (2000). Recursive Parameter Estimation for Nonlinear Continuous Time Systems through Sensitivity Model Based Adaptive Filters. PhD thesis. University of Bochum.

Davies, R. (2004). Newmat c++ matrix library. In: *www.robertnz.net/nm_intro.htm*.

Hovland, G., S. Hanssen, E. Gallestey, S. Moberg, T. Brogårdh, S. Gunnarsson and M. Isaksson (2002). Nonlinear identification of backlash in robot transmissions. In: *Proceedings of the 33rd International Symposium on Robotics (ISR2002)*. Stockholm.

Julier, S.J. and J.K. Uhlmann (1997). A new extension of the kalman filter to nonlinear systems. In: *Proceedings of AeroSens: The 11th Int. Symp. on Aerospace/Defense Sensing, Simulation and Control*.

Julier, S.J., J.K. Uhlmann and H.F. Durrant-Whyte (2000). A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*.

Lagerberg, A. and B. Egardt (2003). Estimation of backlash with application to automotive powertrains. In: *Proc. of the 42nd IEEE Conference on Decision and Control*.

Ljung, L. (1979). Asymptotic behaviour of the extended kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*.

Ljung, L. (1999). *System Identification: Theory for the User*. Prentice Hall.

Ljung, L. and T. Söderström (1993). *Theory and Practice of Recursive Identification*. MIT Press series in signal processing, optimization, and control ; 4.

Skogestad, S. and I. Postlethwaite (1996). *Multivariable Feedback Control: Analsysis and Design*. John Wiley & Sons. New York.

Wan, E.A. and R. van der Merwe (2000). The unscented kalman filter for nonlinear estimation. In: *Proceedings of Symposium 2000 on Adaptive Systems for Signal Processing, Communication and Control (AS-SPCC)*. Alberta, Canada.

Wan, E.A. and R. van der Merwe (2001*a*). *Kalman Filtering and Neural Networks*. Chap. 7. John Wiley.

Wan, E.A. and R. van der Merwe (2001*b*). The square-root unscented kalman filter for state and parameter estimation. In: *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing)*. Salt Lake City.