# State estimation IS the real challenge in NMPC

Steinar Kolås[†], Bjarne Foss[‡] and Tor Steinar Schei[⋆]

[†] *Hydro Aluminium ASA, Technology and Operational Support, N-6884 Øvre Årdal,* steinar.kolas@hydro.com
[‡] *Department of Engineering Cybernetics, NTNU, N-7491 Trondheim, Norway,* bjarne.foss@itk.ntnu.no
[⋆] *Cybernetica AS, N-7038 Trondheim, Norway,* tor.steinar.schei@cybernetica.no

**Keywords** : Unscented, Divided Difference, Kalman filter, Nonlinear state estimation, Constrained nonlinear systems, Noise modeling.

**Abstract** : Nonlinear model predictive control (NMPC) opens for the use of MPC in more demanding applications than has normally been the case for linear MPC. Hence, NMPC is gaining wider acceptance as a technology for challenging control problems within the process industries. The scope of this paper is to discuss an important challenge when applying NMPC, namely the state estimation problem. Our experience shows that this is the critical challenge in most cases. To ensure a sound level of credibility we first present some applications and typical control challenges in some detail. Thereafter, we discuss approaches to get to grips with the state estimation problem in a sound manner focussing on two issues: state estimation methods and noise modeling. The importance of applying higher order filters as an alternative to the extended Kalman Filter (EKF) in demanding applications is highligthed. To overcome limitations in the EKF, several approaches and alternatives have been suggested. We will focus on some of the differential-free algorithms - the Divided Difference (or Central Difference) Kalman filter (DDKF/CDKF) and the Unscented Kalman filter (UKF), and our experience with these approaches. Recursive state estimation algorithms usually assumes that uncertainty enters through additive white noise sources. We discuss noise modeling based on a hypothesis that it is important to model noise correctly. In practice this implies a critical view on the dominating 'additive noise paradigm' as a means to model uncertainty.

Int. Workshop on Assessment and Future Directions of NMPC
Pavia, Italy, September 5-9, 2008

# 1 Introduction

Nonlinear model predictive control (NMPC) opens for the use of MPC in more demanding applications than has normally been the case for linear MPC. In particular NMPC lends itself to nonlinear systems which exhibit large variations in operating conditions and which are critically dependent on the use of a dynamic nonlinear model to gain sufficient performance. Hence, NMPC is gaining wider acceptance as a technology for challenging control problems within the process industries. NMPC is not a well defined term in the sense that NMPC may be used for controllers ranging from a slight variation of linear MPC to the online solution of a constrained nonlinear optimization problem. One example of a slight modification to account for nonlinearity is the use of multiple linear models in such a way that the current working point defines which model should be active at a given time instant. Hence, the QP-problem frequently encountered in linear MPC will change as the active model changes. In our context NMPC shall mean the use of a nonlinear mechanistic model, state estimation, and the solution of an online constrained nonlinear optimization problem.

The scope of this paper is to discuss an important challenge when applying NMPC, namely the state estimation problem. Our experience shows that this is the critical challenge in most cases. To ensure a sound level of credibility we first present some applications and typical control challenges in some detail. Thereafter, we discuss approaches to get to grips with the state estimation problem in a sound manner focussing on two issues: state estimation methods and noise modeling. The importance of applying higher order filters as an alternative to the extended Kalman Filter (EKF) in demanding applications is highligthed. To overcome limitations in the EKF, several approaches and alternatives have been suggested. We will focus on some of the differential-free algorithms - the Divided Difference (or Central Difference) Kalman filter (DDKF/CDKF) and the Unscented Kalman filter (UKF), and our experience with these approaches.

Recursive state estimation algorithms usually assumes that uncertainty enters through additive white noise sources. Further, unknown and time-varying parameters are often modelled similarly by augmenting the states with a parameter vector. Finally, initial model uncertainty is reflected through the choice of the initial covariance matrices for the states and parameters. We discuss noise modeling based on a hypothesis that it is important to model noise correctly. In practice this implies a critical view on the dominating 'additive noise paradigm' as a means to model uncertainty.

The paper starts with a motivating example before some estimation algorithms and noise models are explored. A discussion and some conclusions end the paper.

# 2 A motivating example - the Hall-Heroult process

The Hall-Heroult process is the dominating process for producing aluminum today ([10]). The fundamentals of the process are to dissolve $Al_2O_3$ in molten cryolite (also known as electrolyte or bath), and electrically reduce complex aluminum containing ions to pure aluminum. The process has strong internal couplings, for instance between the mass and energy balance through the side ledge. The coupled mass and energy balance combined with nonlinear process characteristics and few measurements, makes the Hall-Heroult process challenging to control ([7], [6], [9] ).

Recently Hydro Aluminium have been active in developing an advanced control structure, by initiating an NMPC project that has resulted in a patent application for NMPC control of the Hall-Heroult process ([15]). An important challenge in an NMPC application is connected to the estimator, in that the complexity and efficiency of the NMPC is closely related to the quality of the estimates produced by the estimator. This is illustrated in Figure 1, where data from one of the early tests of NMPC in closed loop control of the Hall-Heroult process is shown.

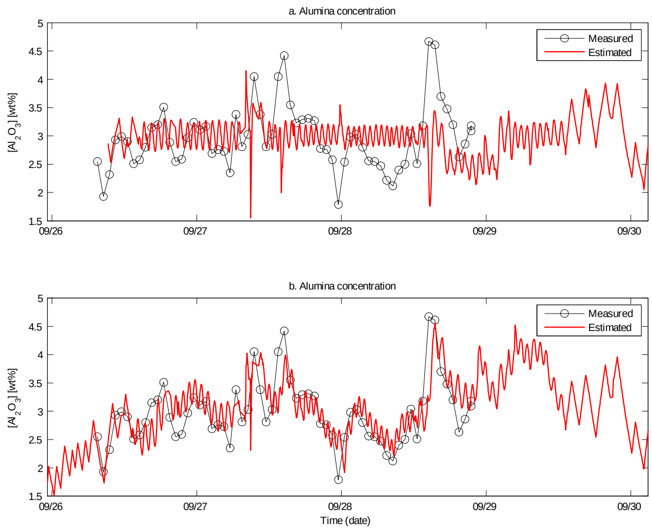


Figure 1: The figure shows measured and estimated alumina concentration for different tuning (a) and (b) of an estimator for the Hall-Heroult process. Note that the measured alumina concentration is not available to the estimator.

Figure 1 clearly illustrates that the performance of the estimator is crucial for the expected performance of the NMPC application. The quality of the estimates may not only be dependent on the accuracy of the model, but also of the estimating method selected and how process knowledge is applied ([17]).

## 3 Nonlinear state estimation

Nonlinear state estimation is a field with broad contributions[1]. It includes nonlinear estimators such as Moving Horizon Estimation, the Particle filter, the Ensemble Kalman filter, the Unscented Kalman filter and the Extended Kalman filter, just to mention some. The Extended Kalman filter (EKF), which was originally proposed by Stanley Schmidth in 1967 ([1]) in order to apply the Kalman filter to nonlinear spacecraft navigation problems, is probably the most

[1] For a review of nonlinear state estimation see [4] and [5].

used method in applied nonlinear state estimation. However, several authors have experienced shortcomings applying EKF to systems with severe nonlinearity and/or constraints (see e.g. [3], [24], [12], [13], [14], [19], [20], [11], [2] to mention some). The shortcomings are related to difficulties in determining the Jacobians, errors introduced by linearization and/or to deal with systems with multimodal or asymmetric probability density functions (pdf). Also, if handling of constraints are unavoidable, the EKF has some limitations in propagating the constraints both through the states and covariance calculations.

In this work we use the reformulated UKF algorithms as described in [16]. When using UKF with constraint, we constrain the sigma points and the updated sigmapoints. The constraint in this work is implemented by projection similar to 'clipping' (for further references on constraints and UKF see [16]).

Also the Divided Difference Kalman filter (DDKF) is briefly considered in this work ([19]). Further, the analysis on how noise enters the covariance is done using a continuos time formulation, and we assume these results also are valid for the discrete time case. Based on this assumption, and for the case of simplicity regarding the theoretical results presented, we investigate the covariance equation found in the continuos EKF ([8])

$$\dot{P} = FP + PF^T + HQH^T - KRK^T \tag{1}$$

where all the elements in (1) may be time varying. Here $K$ is the Kalman gain typically found in he Kalman filter, $F$ and $H$ denote the Jacobians found from an appropriate system model. $Q$ is the assumed process noise covariance and $R$ is the assumed measurement noise covariance. Especially the term $HQH^T$ in (1) will have our attention in this work, since this is the term that determines how the system noise is injected into the covariance calculation. A more detailed discussion about noise and modeling is to be found in [17].

# 4 Estimation in constrained nonlinear systems

## 4.1 Investigated case - 2 state CSTR

The results from the 'motivating example' motivated further studies of the UKF approach. In [16] a broad overview of different UKF algorithms is given, and extensions to the ensemble of UKF algorithms are suggested, discussing in particular the use of constraints. In the sections to come, the performance of the constrained approach is compared with EKF, a selection of UKF algorithms and some DDKF algorithms applied to a 2 state nonlinear CSTR example process with a multi-modal probability density functions. This is a very challenging state estimation problem.

### 4.1.1 Case description

Consider the gas-phase, reversible reaction [11]

$$2A \xrightarrow{k_r} B \quad k_r = 0.16 \tag{2}$$

with stoichiometric matrix

$$s = \begin{bmatrix} s_1 & s_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \end{bmatrix} \tag{3}$$

and reaction rate

$$r = k_r P_A^2 \tag{4}$$

The state and output vectors are defined as

$$x = \begin{bmatrix} P_A \\ P_B \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad y = \begin{bmatrix} 1 & 1 \end{bmatrix} x \tag{5}$$

where $P_A$ and $P_B$ are the partial pressures. It is assumed that the ideal gas law holds (high temperature, low pressure), and that the reaction occurs in a well-mixed, isothermal batch reactor.

The model used is the discrete analytical solution as found in [11]. Further it is assumed that the system experiences Gaussian noise both in the states and in the outputs, respectively $v_k \sim N(0, Q_k)$ and $w_k \sim N(0, R_k)$.

The parameters used for this system are $\Delta t = t_{k+1} - t_k = 0.1$, $P_0 = diag^2 \begin{bmatrix} 6^2 & 6^2 \end{bmatrix}$, $k_r = 0.16$, $x_0 = \begin{bmatrix} 3 & 1 \end{bmatrix}^T$, $x_0 = \begin{bmatrix} 3 & 1 \end{bmatrix}^T$, $\hat{x}_0 = \begin{bmatrix} 0.1 & 4.5 \end{bmatrix}^T$, $Q = diag \begin{bmatrix} 0.001^2 & 0.001^2 \end{bmatrix}$ and $R = 0.1^2$. Note that the initial guess for the estimator ($\hat{x}_0$), is very poor. This simple example is used by several authors in order to investigate estimator performance (see [11], [24], [21], [14]). The reason why this problem is interesting is that the estimator may experience a multimodal probability density function, which may lead to unphysical estimates.

## 4.2 Simulation results

In the following chapters we will investigate some estimation algorithms applied to the 2-state CSTR case. Note that all the parameters are as described in the case description above for all the presented algorithms. In the presented results also the noise sequences are identical in all simulations, except for the DDKF. We have chosen to be true to the source of these examples [11] and have used the same parameters to achieve comparable results.

### 4.2.1 EKF

The EKF algorithm with numerically derived Jacobians is investigated. Figure 2 shows the results of the simulation using unconstrained EKF.

The unconstrained EKF fails to converge to the true states within the given time frame[3]. These results are in agreement with the results of [11] and [14]. The reason why the EKF fails is that while the negative pressure is unphysical, the unconstrained estimator allows the estimate to enter regions where the partial pressure may be negative. By using EKF with constraints, where the constraints are implemented by clipping the corrected state estimate such that $\hat{x}_k \geq 0$, again the constrained EKF fails in converging to the true states. These results are in agreement with the results of [11] and [14]. By clipping the state $\hat{x}_k$, it is restricted to a valid physical region, but the knowledge about the constraints is not propagated into the covariance, and hence the accuracy of the approximated covariance matrix $P_{x_k}$ is questionable.

[2] $diag()$ is an operator creating an $nxn$ matrix with the given elements on the diagonal and 0 in all other entries.

[3] Actually the EKF will converge very slovly, but one need to run the simulation approx. 1000 samples.

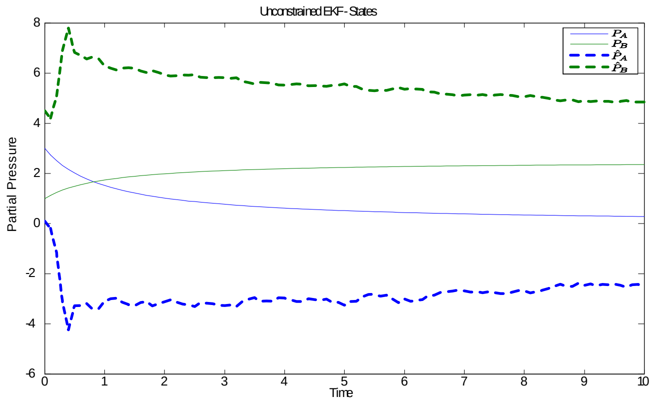


Figure 2: Unconstrained EKF.

### 4.2.2 UKF

The unconstrained UKF and the reformulated UKF, with the same tuning as the EKF, is applied to the test case.

As Figure 3a shows, the unconstrained UKF fails to converge to the true states within the given time frame. These results are in agreement with the results of [14]. The unconstrained UKF as well as the EKF has to deal with multiple optima, and we believe this is the reason why also the unconstrained UKF algorithms suffer poor performance on this case. All the investigated UKF algorithms converge to the true state when constraining the sigma points $\chi^x_{k-1} \geq 0$. This is shown in Figure 3b for one realization of the reformulated UKF, applied to the non-augmented UKF algorithm proposed by [22], see also [16]. The performance in this case is excellent.

### 4.2.3 Divided Difference KF

The Divided Difference Kalman filter (DDKF) is based on polynomial approximations of the nonlinear transformations obtained with a multidimensional extension of Stirling's interpolation formula. DD1 is based on a first order approximation and DD2 is based on a second order approximation [19]. As the UKF algorithms, the DDKF algorithms are Jacobian free algorithms and the mean and covariance are in principal calculated in the same manner as in the UKF - by some weighted sums. In [25] it is shown that the DD2 KF has a slightly more accurate covariance estimate than the UKF based on the scaled unscented transform (SUT), but for all practical purposes there are no difference in estimation performance between the SUT based UKF and the DD2 KF [25].

DD1 and DD2 KF are tested on the '2 state CSTR' case. The noise is assumed Gaussian. Note that the noise sequences used in the simulations for the DD1 KF and DD2 KF are not identical, and are also different than the noise sequence used in the previous section discussing EKF and UKF. The initial values and system parameters are the same as for the EKF and UKF simulations previously discussed. By constraining the corrected state estimate

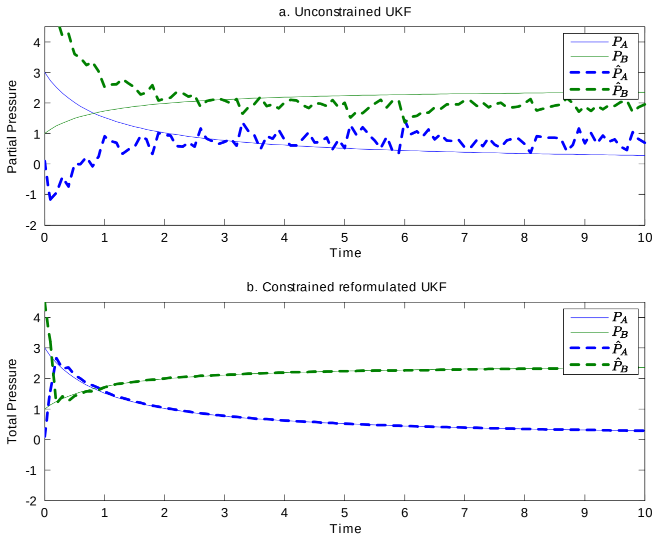


Figure 3: The figure shows a. the results of using the unconstrained fully augmented UKF as by [25] and b. the results of using the constrained reformulated UKF as proposed in [16]. Constraints used is $CC_1$ and $CC_7$ and the non-augmented UKF.

such that $\hat{x}_k \geq 0$, we get the results as given in Figure 4.

As Figure 4 shows, DD1 KF outperforms the EKF and converges to the true states, but not as fast as the constrained UKF algorithms discussed above. Also DD2 KF outperforms EKF and converge to the true states. The convergence speed is similar to the constrained non-augmented reformulated UKF algorithm discussed above.

## 5 Noise modeling concepts

As claimed earlier, the quality of the estimates may not only be dependent on the accuracy of the model and of the estimating method selected, but also of how process knowledge is applied. In our view, noise modeling has attracted more attention in the system identification literature (see e.g. [18]) than in the literature on identification and estimation for physics based models. One example is how white noise may enter an input-output model in different ways. Studying noise modeling based on a hypothesis that it is important to model noise correctly, implies in practice a critical view on the dominating 'additive noise paradigm' as a means to model uncertainty. The 'additive noise paradigm' dominates textbooks and papers on recursive state estimation, i.e. Kalman filter type algorithms like the EKF and UKF (see e.g. [8], [22], [23]) even though uncertainty may enter a system in many different ways. The additive noise model structure is obviously reasonable in many applications. In others, however, this is not the case. One example are processes where control input uncertainty dominates, and where this noise depends on the value of the control

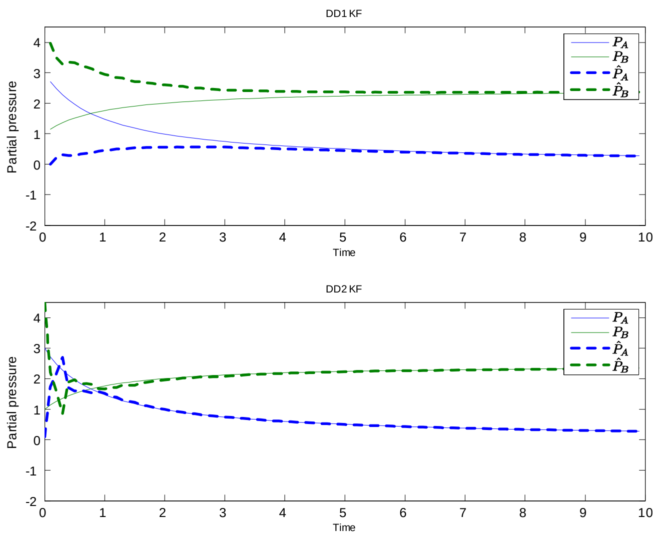


Figure 4: The simulation results using the DD1 KF (upper most subplot) and the DD2 KF (lower most subplot).

input itself. It may for instance increase proportionally with the control input. A fruitful way to view noise modeling is to view this as a direct extension of the process of developing a model. We assume dynamic models which are developed using physical insight and process data, i.e. physics-based models. Having established and possibly validated such a model, it is at least in principle possible to quantify uncertainty. This may include uncertainty in initial conditions, and in certain time-varying states and parameters, control inputs and measurements. Further, it may be possible to describe how noise enters the system, i.e. to structurally model how uncertainty affects the model.

Recursive state estimation algorithms usually assume that uncertainty enters through additive white noise sources. Further, unknown and time-varying parameters are often modeled similarly by augmenting the states with a parameter vector. Finally, initial model uncertainty is reflected through the choice of the initial covariance matrices for the states and parameters. In the sections to follow we investigate the effect of different system noise modeling methods.

## 5.1 Method 1 - Additive noise

Consider the gas-phase, reversible reaction as given by (2)-(5). Further we assume the reaction parameter $k_r$ varies and is modelled as colored noise, and is estimated. The state and output vectors for the estimator are defined as

$$\hat{x} = \begin{bmatrix} \hat{P}_A \\ \hat{P}_B \\ \hat{k}_r \end{bmatrix} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix}, \qquad \hat{y} = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \hat{x} \tag{6}$$

and extended with control inputs ($u$). Described the traditional way by assuming additive noise on the system we get:

$$\dot{\hat{x}}_1 = -2\hat{r} + u_1 + v_1 \tag{7}$$

$$\dot{\hat{x}}_2 = \hat{r} + u_2 + v_2 \tag{8}$$

$$\dot{\hat{x}}_3 = v_{k_r} \tag{9}$$

where the state vector $\hat{x}$ is given by (6) and the reaction rate $\hat{r}$ is given by

$$\hat{r} = \hat{k}_r \hat{x}_1^2 = \hat{x}_3 \hat{x}_1^2 \tag{10}$$

The noise vector is given by

$$v = \begin{bmatrix} v_1 & v_2 & v_{k_r} \end{bmatrix}^T \tag{11}$$

and the parameters by

$$\hat{\theta} = \hat{k}_r \tag{12}$$

Assume that the system (2)-(5) is considered as a semi-batch system in that the species $B$ is removed and that the species $A$ is refilled when a certain level of $A$ is reached resulting in the control input described by (13)

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \left\{ \begin{array}{c} 4 - \hat{x}_1) \\ -\hat{x}_2 \end{array} \right\} \hat{x}_1 \leq 0.2 \\ \\ \left\{ \begin{array}{c} 0 \\ 0 \end{array} \right\} \hat{x}_1 > 0.2 \end{bmatrix} \tag{13}$$

The Jacobian $H$ is given by the identity matrix $I$, and hence $HQH^T$ in (1) is given by

$$HQH^T = Q = diag \begin{bmatrix} v_1^2 & v_2^2 & v_{k_r}^2 \end{bmatrix} \tag{14}$$

It is common practice to look at the state covariance as if it reflects the confidence in the system states $\hat{x}$, and that quantitatively a relatively high value means quite uncertain estimates and vice versa ([23, p. 326]). A challenge using the proposed noise formulation might be that if the covariance has settled to a relative low value (i.e. certain state estimates) when the control inputs is applied, i.e. removing $B$ and refilling $A$, this may introduce large errors in the estimates. Also worth noticing is that with this formulation there is no information about the control inputs in the covariance equation given by (1).

## 5.2 Method 2 - Noise in control inputs

We now consider that there is some uncertainty related to the control inputs, and that the uncertainty in the inputs can be expressed as a relative uncertainty. That is:

$$\dot{\hat{x}}_1 = -2\hat{r} + u_1(1 + v_{u_1}) \tag{15}$$

$$\dot{\hat{x}}_2 = \hat{r} + u_2(1 + v_{u_2}) \tag{16}$$

$$\dot{\hat{x}}_3 = v_{k_r} \tag{17}$$

where the state vector $\hat{x}$ is given by (6) and the reaction rate $\hat{r}$ is given by

$$\hat{r} = \hat{k}_r \hat{x}_1^2 = \hat{x}_3 \hat{x}_1^2 \tag{18}$$

The noise vector is given by

$$v = [\ v_{u_1} \quad v_{u_2} \quad v_{k_r}\ ]^T \tag{19}$$

and the parameters by (12). The Jacobian $H$ becomes

$$H = diag \left[\ u_1 \quad u_2 \quad 1\ \right] \tag{20}$$

Assuming the noise is described by

$$Q = diag \left[\ v_{u_1}^2 \quad v_{u_2}^2 \quad v_{k_r}^2\ \right] \tag{21}$$

the term $HQH^T$ in (1) becomes

$$HQH^T = diag \left[\ u_1^2 v_{u_1}^2 \quad u_2^2 v_{u_2}^2 \quad v_{k_r}^2\ \right] \tag{22}$$

Consider again that the covariance reflects the uncertainty of the system states $\hat{x}$. In the case when the control inputs are applied, large errors in the estimates may be introduced. This is reflected in the proposed formulation (22) by the injection of the input uncertainty into the covariance function.

## 5.3 Method 3 - Noise in auxiliary variables

In this case the noise enters the system through the reaction rate $r$, instead of directly on the system states. That is:

$$\dot{\hat{x}}_1 = -2\hat{r} \tag{23}$$

$$\dot{\hat{x}}_2 = \hat{r} \tag{24}$$

$$\dot{\hat{x}}_3 = v_{k_r} \tag{25}$$

where the state vector $\hat{x}$ is given by (6) and the reaction rate $\hat{r}$ is given by

$$\hat{r} = \hat{k}_r \hat{x}_1^2 + v_r = \hat{x}_3 \hat{x}_1^2 + v_r \tag{26}$$

The noise vector is given by

$$v = [\ v_r \quad v_{k_r}\ ]^T \tag{27}$$

and the parameters by (12). The system becomes

$$\dot{\hat{x}}_1 = -2\hat{k}_r \hat{x}_1^2 - 2v_r \tag{28}$$

$$\dot{\hat{x}}_2 = \hat{k}_r \hat{x}_1^2 + v_r \tag{29}$$

$$\dot{\hat{x}}_3 = v_{k_r} \tag{30}$$

and the Jacobian $H$

$$H = \begin{bmatrix} -2 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{31}$$

Assuming the noise is described by

$$Q = diag \begin{bmatrix} v_r^2 & v_{k_r}^2 \end{bmatrix} \tag{32}$$

the term $HQH^T$ in (1) becomes

$$HQH^T = \begin{bmatrix} 4v_r^2 & -2v_r^2 & 0 \\ -2v_r^2 & v_r^2 & 0 \\ 0 & 0 & v_{k_r}^2 \end{bmatrix} \tag{33}$$

That is, by applying noise on the auxiliary variable $\hat{r}$, correlation is naturally introduced in the covariance calculation. This could also be seen as an alternative to off-diagonal tuning of $Q$ in (14), in that the correlation enters the system naturally and correctly scaled in the off-diagonal elements.

## 5.4 Simulation results

### 5.4.1 2 state CSTR

We investigate the case as described by (2)-(6) using the UKF with constraint handling as in Chapter 5.2. Regarding the discrete system representation, the analytical model as given by [11] is used. Further, the parameters for this system, if not otherwise noted, are $\Delta t = t_{k+1} - t_k = 0.1$, $P_0 = diag \begin{bmatrix} 6^2 & 6^2 & 0.015^2 \end{bmatrix}$, $k_r = 0.16$, $x_0 = \begin{bmatrix} 3 & 1 \end{bmatrix}^T$, $\hat{x}_0 = \begin{bmatrix} 0.1 & 4.5 & 0.9k_r \end{bmatrix}^T$. Note that the initial guess ($\hat{x}_0$) for the estimator is poor. The reaction parameter $k_r$ is constant, but has a wrong initial estimate. The following constraints are applied to the UKF sigma points

$$\begin{array}{r} Lower\ bounds:\ [0, 0, 0.1]^T \\ Upper\ bounds: [\infty,\ \infty,\ , 0.18]^T \end{array} \tag{34}$$

The state estimator used in this work is the fully augmented UKF with reformulation of the correction steps and the use of constraints as presented in [16].

**Method 1** The noise is modeled as in Method 1, the estimator constraints as (34) and the estimator tuning as

$$\begin{array}{c} v = \begin{bmatrix} v_1 & v_2 & v_{k_r} \end{bmatrix}^T = \begin{bmatrix} 10^{-9} & 10^{-9} & 10^{-4} \end{bmatrix}^T \\ w = 0.002 \end{array} \tag{35}$$

In the case when the true process experiences no noise, we get the results as shown in Figure 5.

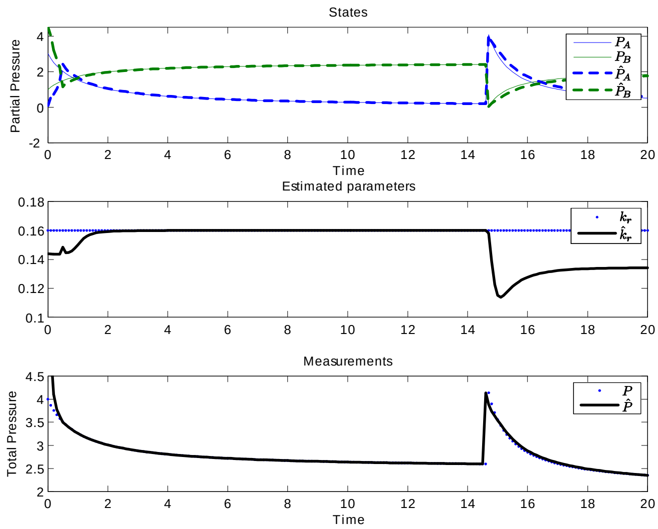


Figure 5: The figure shows the true and estimated states, the true and estimated reaction 'constant' $k_r$ and the true and estimated output using Method 1. As the figure shows, the estimates of $k_r$ is disturbed in the transient periode caused by the control input exitation.

As Figure 5 shows, the bad initial guess is handled very well, and the state estimate and output estimate is generally acceptable. However the parameter estimate during the transient period at approximately $t = 15$ is generally bad due to the control input excitation. Inspection of the time trace of the covariance is as expected, i.e. the control input excitation is not reflected in the covariance.

By modeling the noise by combining Method 1 and 2, and by that introduce uncertainty also in the control inputs, the system equations becomes

$$\dot{\hat{x}}_1 = -2\hat{r} + u_1(1 + v_{u_1}) + v_1 \tag{36}$$

$$\dot{\hat{x}}_2 = \hat{r} + u_2(1 + v_{u_2}) + v_2 \tag{37}$$

$$\dot{\hat{x}}_3 = v_{k_r} \tag{38}$$

where the state vector $\hat{x}$ is given by (6) and the reaction rate $\hat{r}$ is given by

$$\hat{r} = \hat{k}_r \hat{x}_1^2 = \hat{x}_3 \hat{x}_1^2 \tag{39}$$

The noise vector is given by $v = \left[ \begin{array}{ccccc} v_1 & v_2 & v_{k_r} & v_{u_1} & v_{u_2} \end{array} \right]^T$. $HQH^T$ in (1) becomes

$$HQH^T = diag \left[ \begin{array}{ccc} v_1^2 + u_1^2 v_{u_1}^2 & v_2^2 + u_2^2 v_{u_2}^2 & v_{k_r}^2 \end{array} \right] \tag{40}$$

With the estimator constraints as by (34), and the estimator tuning as

$$\begin{gathered} v = \left[ \begin{array}{ccccc} v_1 & v_2 & v_{k_r} & v_{u_1} & v_{u_2} \end{array} \right]^T = \left[ \begin{array}{ccccc} 10^{-9} & 10^{-9} & 10^{-4} & 1 & 1 \end{array} \right]^T \\ w = 0.002 \end{gathered} \tag{41}$$

the results for the case when the simulator experience no noise, is as shown in Figure 6.

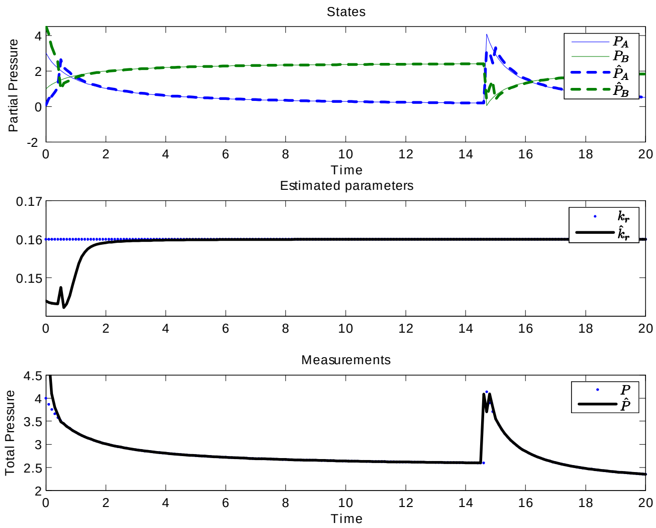


Figure 6: The figure shows the true and estimated states, the true and estimated reaction 'constant' $k_r$ and the true and estimated output. As the figure shows the estimated $k_r$ is not disturbed in the transient periode caused by the control input exitation.

As Figure 6 shows, the estimate of the parameter $k_r$ does not suffer from the control input excitation. However the 'cost' is some loss off accuracy after the control input excitation. Inspection of the time trace of the covariance shows that the control input excitation at approximately $t = 15$ is reflected in the covariance elements $(P_{x_k}(1,1),\ P_{x_k}(1,2),\ P_{x_k}(2,1),\ P_{x_k}(2,2))$. That is, by applying the idea that the covariance should reflect the uncertainty in the states around the control input excitation, good results are achieved by combining the concept in Method 1 and 2 (see [17] for further references).

# 6 Discussion

This paper centres on performance of different nonlinear state estimators and on noise modeling . Three different state estimators have been tested on a simple and challenging problem showing quite different performance. This is not surprising. It does however indicate that the use of higher order filters may be required for demanding NMPC applications. Further, the use of constraint handling can be critical to obtain good estimator performance. This study is significantly extended in [16]. The results in this reference align with conclusions herein.

Noise modeling is significant for estimator performance. Again a simple example is used to illustrate this fact. A recent study [17] expands on the results provided herein. In particular the conclusions coincide with the observations in this paper.

The results in this paper are consistent with our experience for NMPC applications in which the choice of nonlinear filter algorithms and the choice of reasonable noise model are critical factors to obtain good NMPC performance.

# 7 Conclusions

Our experience shows that the state estimation problem is indeed the critical challenge in most NMPC applications. Furthermore, practice reveals that the use of higher order filters like the UKF is important to obtain good performance in challenging applications. Finally, alternative noise models to the dominating 'additive noise paradigm' has a positive impact on estimation performance in some cases.

# 8 Acknowledgements

The financial support of Hydro Aluminium ASA is gratefully acknowledged.

# References

[1] Bellatoni, J. and Dodge, K. (1967). A square root formulation of the Kalman Schmidt filter. *AIAA Journal*, 5:1309–1314.

[2] Bizup, D. F. and Brown, D. E. (2003). The over-extended Kalman filter - dont use it! *Proceedings of the Sixth International Conference of*, 1:40–46.

[3] Chen, W., Lang, L., Bakshi, B. R., Goel, P. K., and Ungarla, S. (2006). Bayesian Estimation via Sequential Monte Carlo Sampling - Constrained Dynamic Systems. *Preprint submitted to Elsevier Science.*

[4] Daum, F. E. (1986). Beyond Kalman filters: practical design of nonlinear filters. *SPIE*, 2561:252–262.

[5] Daum, F. E. (2005). Nonlinear Filters:Beyond the Kalman filters. *IEEE AE system magazine*, 20(8):57–69.

[6] Drengstig, T., Ljungquist, D., and Foss, B. (1998). On the AlF3 and temperature control of an aluminumelectrolysis cell. *IEEE Transactions on Control Systems Technology*, 6:157–171.

[7] Foss, B. and Schei, T. S. (2005). Putting Nonlinear Model Predictive Control into Use. *Workshop on Nonlinear Model Predictive Control*, Freudenstadt(Germany).

[8] Gelb, A. (1974). Applied Optimal Estimation. *The M.I.T. Press*, ISBN 0-262-20027-9.

[9] Gran, E. (1980). A Multi-Variable Control in Aluminum Reduction Cells. *Modeling Identification and Control*, 1(4):247–258.

[10] Grjotheim, K. and Kvande, H. (1993). Introduction to aluminium electrolysis. *Aluminum-Verlag.*

[11] Hasseltine, E. L. and Rawlings, J. B. (2003). A Critical evaluation of Extended Kalman Filtering and Moving Hoizon Estimation. *TWMCC, Technial report number 2002-03.*

[12] Julier, S. J. and Uhlmann, J. K. (1994). A general method for approximating nonlinear transformation of probability distributions. *Technical report, RRG, Dept. of Engineering Science, University of Oxford*, Can be downloaded from http://www.robots.ox.ac.uk/~siju/work/publications/letter size/Unscented.zip, August 1994.

[13] Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. (1995). A new approach for filtering nonlinear system. *Proceedings of the American Control Conference*, pages 1628–1632.

[14] Kandepu, R., Foss, B., and Imsland, L. (2008). Applying the Unscented Kalman Filter for Nonlinear State Estimation. *Journal of Process Control*, 18:753–768.

[15] Kolås, S. (2007). Method and means for controlling an electrolysis cell. *Norsk Hydro Patent Application*, NO 20075933.

[16] Kolås, S., Foss, B., and Schei, T. S. (2008a). Constrained nonlinear state estimation based on the UKF approach. *to Computers and Chemichal Engineering*, Submitted March 2008.

[17] Kolås, S., Foss, B., and Schei, T. S. (2008b). Noise modeling concepts in Nonlinear State Estimation. *Journal of Process Control*, Submitted April 2008.

[18] Ljung, L. (1987). System Identification. Theory for the User. *Prentice-Hall*, Englewodd Cliffs(NJ).

[19] Nørgaard, M., Poulsen, N. K., and Ravn, O. (2000). New developments in state estimation for nonlinear systems. *Automatica*, 36:1627–1638.

[20] Rao, C. V. (2000). Moving Horizon Strategies for the Constrained Monitoring and Control of Nonlinear Discrete-Time Systems. *PhD thesis (Chemical Engineering)*, Univeristy of Wisconsin-Madison.

[21] Rawlings, J. B. and Bakshi, B. R. (2006). Particle filter and moving horizon estimation. *Computers and Chemichal Engineering*, 30:1529–1541.

[22] Simon, D. (2006). *Optimal State Estimation.* Wiley, ISBN-10 0-471-70858-5.

[23] Søderstrøm, T. and Stoica, P. (1989). System identification. *Prentice Hall.*

[24] Vachhani, P., Narasimhan, S., and Rengaswamy, R. (2006). Robust and reliable estimation via Unscented Recursive Nonlinear Dunamic Data Reconsiliation. *Jounal of Process Control*, 16:1075–1086.

[25] van der Merwe, R. (2004). Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models. *PhD dissertation submitted to the faculty of the OGI School of Science and Engineering at Oregon Health and Science University.*