Noise modeling concepts in nonlinear state estimation

S. Kolås [a,b,*], B.A. Foss [a], T.S. Schei [c]

[a] Department of Engineering Cybernetics, NTNU, N-7491 Trondheim, Norway
[b] Hydro Aluminium AS Technology and Operational Support, N-6884 Øvre Årdal, Norway
[c] Cybernetica AS, N-7038 Trondheim, Norway

ARTICLE INFO

ABSTRACT

It is quite common to assume that uncertainty enters through additive white noise sources when using recursive state estimation algorithms. Also unknown and time-varying parameters are often modeled similarly by augmenting the states with a parameter vector. Further, it is common to reflect initial model uncertainty through the choice of the initial covariance matrices for the states and parameters.

In this paper we study 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. Alternative concepts of modeling the noise are investigated, and it is shown that modeling noise by introducing it in the system auxiliary variables and control inputs may have a positive impact on estimation performance.

1. Introduction

Recursive state estimation algorithms usually assumes 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.

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. [25,26,6]) 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 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.

It should be added that noise modeling has attracted more attention in the system identification literature (see e.g. [17]). One example is how white noise may enter an input–output model in different ways.

In this paper alternative noise modeling concepts are investigated, and it is shown that introducing noise in auxiliary variables and control inputs may have a positive impact on estimation performance.

The paper is organized as follows: First noise modeling concepts are presented. Then follows a simulation study based on available literature examples and to an industrial process. Finally, some discussion and conclusions ends the paper.

2. System description

In this work we address the general continuous nonlinear system given by the state space formulation

$$\dot{x}(t) = f_c(x(t), \theta(t), u(t), v(t)) \tag{1}$$
$$x(0) - \text{given}$$
$$y(t) = h(x(t), w(t)) \tag{2}$$

where $x(t)$ denotes the system states, $\theta(t)$ denotes the system parameters, $y(t)$ the output model (or measurements), $u(t)$ control inputs, $v(t)$ system state noise input (stationary or time-varying

stochastic), $w(t)$ output (measurement) noise inputs (stationary or time-varying stochastic), and $f_c(\cdot)$ and $h(\cdot)$ are nonlinear Lipschitz continuous functions. All variables are vectors with appropriate dimensions.

Further we address the general discrete time nonlinear system with constant sampling time (denoted by k) given by the state space formulation

$$x_k = f(x_{k-1}, \theta_{k-1}, u_{k-1}, v_{k-1}) \tag{3}$$
$$x_0 - \text{given}$$
$$y_{k-1} = h(x_{k-1}, w_{k-1}) \tag{4}$$

where x_k denotes the system states, θ_k denotes the system parameters, y_k the output model (or measurements), u_k control inputs, v_k system state noise input (stationary or time-varying stochastic), w_k output (measurement) noise inputs (stationary or time-varying stochastic), $f(\cdot)$[1] and $h(\cdot)$ are nonlinear functions.

All variables are vectors with appropriate dimensions. The model used in the estimators and estimated values are denoted with a $\wedge$, typically $\hat{x}_k, \hat{y}_k$, etc.

We have chosen to present the analytical work using the continuous formulation, while all the simulations and estimations are done in discrete time. We assume that the analytical results valid for the continuous time case also are valid for the discrete time case.

2.1. State estimation

In 1960 Rudolph Kalman presented a recursive state estimation method, which has become known as the Kalman filter [14]. The important contribution was the fact that a recursive algorithm could be used to accurately compute the first and second order moments (mean and covariance) of a linear system corrupted by Gaussian white noise on its system and output models. Kalman's simple assumptions were that the system random variables could be consistently estimated by sequentially updating the first and second order moments, and that the estimator was on the linear form

$$\hat{x}_k = \hat{x}_k^- + K_k \cdot e_k \tag{5}$$

The Kalman filter consists of two parts; a forward prediction part and a correction part. The prediction part computes a predicted (a prior) estimate of the first and second order moments at time k given information up until $k-1$ denoted by $\hat{x}_k^-$ and $P_{x_k}^-$. The correction part computes the corrected (posterior) estimates $\hat{x}_k$ and P_{x_k} using available data at time k. In (5) the predicted estimate ($\hat{x}_k^-$) is updated by a linear gain (K_k) times an error (e_k). The error part is the deviation between an output value(s) (y_k) and the estimate of the said output(s) ($\hat{y}_k$), and hence (5) can be formulated as

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - \hat{y}_k) \tag{6}$$

where the linear gain (K_k), or the so-called Kalman gain, is given by

$$K_k = P_{x_k y_k} P_{y_k y_k}^{-1} \tag{7}$$

where $P_{x_k y_k}$ and $P_{y_k y_k}$ is the cross covariance and output covariance, respectively.

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 used method in applied nonlinear state estimation. However, several authors have experienced shortcomings

applying EKF to systems with severe nonlinearity and/or constraints (see for example [11,12,23,18,20,9,2,3,27,15] 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.

The Unscented Kalman[2] filter, proposed by Julier and Uhlmann [11],[3] apply almost the same computational structure as the EKF, but is based on the intuition that '*it is easier to approximate a probability distribution than a nonlinear function*' [13,10]). By using the Unscented transform to compute the mean and covariance, the Unscented Kalman filter avoids the need to use Jacobians in the algorithm. The probability distribution is approximated by a set of deterministic points which captures the mean and covariance of the distribution. These points, called sigma points, are then processed through the nonlinear model of the system, producing a set of propagated sigma points. By choosing appropriate weights, the weighted average and the weighted outer product of the transformed points gives the mean (for example $\hat{x}_k^-$) and covariance (for example $P_{x_k}^-$) of the transformed distribution. That is, assume there exist a sigma point set χ_{k-1}^x with $2n$ sigma points, a set of appropriate weights W_i^x and W_i^c, and a nonlinear function (or transformation)

$$x_k = f(x_{k-1}) \tag{8}$$

Passing the vector of sigma points $\chi_{k-1,i}^x$ through the nonlinear function results in the propagated sigma points

$$\chi_{k,i}^x = f(\chi_{k-1,i}^x) \tag{9}$$

and the approximated mean $\hat{x}_k$ and the covariance P_{x_k} could be calculated as

$$\hat{x}_k = \sum_{i=1}^{2n} W_i^x \cdot \chi_{k,i}^x \tag{10}$$

$$P_{x_k} = \sum_{i=1}^{2n} W_i^c (\chi_{k,i}^x - \hat{x}_k)(\chi_{k,i}^x - \hat{x}_k)_k^T \tag{11}$$

In this paper the analysis is done using the continuous 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 continuous EKF [6]

$$\dot{P} = FP + PF^T + GQG^T - KRK^T \tag{12}$$

where all the elements in (12) may be time-varying. It is common practice to look at the covariance (P) as if it reflects the confidence in the system states $\hat{x}$, and that quantitatively a relative high value means quite uncertain and that a relative small value means certain state estimates [26, p. 326]. In (12) F and G denote the Jacobians found by $F = \frac{\partial}{\partial x} f_c(x, \theta, u)|_{x=\hat{x}, \theta=\hat{\theta}, u, v}, G = \frac{\partial}{\partial v} f_c(x, \theta, u, v)|_{v, x, \theta, u}$. Especially the term GQG^T in (12) will have our attention in this work, since this is the term that determines how the uncertainty due to process noise is injected into (12) [6, p. 122]. Further, we have the assumed process noise covariance matrix $E\{v(t)v^T(\tau)\} = Q(t)\delta(t - \tau)$ and the assumed measurement noise covariance matrix $E\{w(t)w^T(\tau)\} = R(t)\delta(t - \tau)$, where $\delta(t - \tau)$ is the Dirac function. Further we have the relationship between the continuous time spectral density Q and the discrete time covariance Q_k by

$$Q = Q_k/\Delta t \tag{13}$$

---

[1] Note that the nonlinear function $f(\cdot)$ is typically a discretized representation of a nonlinear continuous model $f_c(\cdot)$. The discretization method used is typically Runge–Kutta 4th order.

[2] In the original work of Julier and Uhlmann [11] the Unscented Kalman filter was named 'The New Filter'.

[3] We refer to this document as the first work of UKF, since it was submitted to IEEE Transactions on Automatic Control in 1994, although not accepted before in 2000.

and R and the discrete measurement covariance matrix R_k by

$$R = R_k \Delta t \tag{14}$$

where Δt denotes the discrete time step (see [25, p. 233] or [6, p. 121]).

3. Noise modeling concepts

As already mentioned in the *Introduction*, the literature concerning modeling and estimation of stochastic systems usually model the uncertainty or noise as additive to the system equations and/or output equations.

In the sections to follow we investigate the effect of different system noise modeling methods in the scope of the Kalman filter approach to estimation. Describing the ideas covered, are in our opinion best done by introducing them by examples on the system introduced below, because we think it makes the ideas clearer than by using a general formulation.

3.1. Method 1 – additive noise

Consider the gas-phase, reversible reaction [9]

$$2A \xrightarrow{k_r} B \tag{15}$$

with stoichiometric matrix

$$s = \begin{bmatrix} s_1 & s_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \end{bmatrix} \tag{16}$$

and reaction rate

$$r = k_r P_A^2 \tag{17}$$

The state and output vectors are defined as

$$x = \begin{bmatrix} P_A \\ P_B \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad y = \begin{bmatrix} 1 & 1 \end{bmatrix} x \tag{18}$$

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. 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

$$\hat{x} = \begin{bmatrix} \hat{P}_A \\ \hat{P}_B \\ \hat{k}_r \end{bmatrix} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix} \hat{y} = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \hat{x} \tag{19}$$

By assuming that the system (15)–(18) is considered and extended with control inputs (u) and by assuming additive noise on the system, we get the model for the process as

$$\dot{x}_1 = -2r + u_1 + v_1 \tag{20}$$
$$\dot{x}_2 = r + u_2 + v_2 \tag{21}$$

Described the traditional way by assuming additive noise on the system, we get the model for the estimator as

$$\dot{\hat{x}}_1 = -2\hat{r} + u_1 + v_1 \tag{22}$$
$$\dot{\hat{x}}_2 = \hat{r} + u_2 + v_2 \tag{23}$$
$$\dot{\hat{x}}_3 = v_{k_r} \tag{24}$$

where the state vector $\hat{x}$ is given by (19) and the reaction rate $\hat{r}$ is given by

$$\hat{r} = \hat{k}_r \hat{x}_1^2 = \hat{x}_3 \hat{x}_1^2 \tag{25}$$

The noise vector is given by

$$v = \begin{bmatrix} v_1 & v_2 & v_{k_r} \end{bmatrix}^T \tag{26}$$

where $v_1 \sim N(0, \bar{v}_1^2)$, $v_2 \sim N(0, \bar{v}_2^2)$ and $v_{k_r} \sim N(0, \bar{v}_{k_r}^2)$. The parameters are given by

$$\hat{\theta} = \hat{k}_r \tag{27}$$

Note that the noise terms above, as well as in similar models in Sections 3.2 and 3.3, are included to show how the noise is modelled. Hence, it does not imply that the noise terms are measured and used as actual inputs to the model.

Assume that the system (15)–(18) 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 (28)[4]

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \begin{Bmatrix} 4 - x_1 \\ -x_2 \end{Bmatrix} & x_1 \leqslant 0.2 \\ \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} & \text{otherwise} \end{bmatrix} \tag{28}$$

The Jacobian G is given by the identity matrix I, and hence GQG^T in (12) becomes

$$GQG^T = \begin{bmatrix} \bar{v}_1^2 & 0 & 0 \\ 0 & \bar{v}_2^2 & 0 \\ 0 & 0 & \bar{v}_{k_r}^2 \end{bmatrix} \tag{29}$$

As mentioned above, it is common practice to look at the covariance (P) as if it reflects the confidence in the system states $\hat{x}$, and that quantitatively a relative high value means quite uncertain and that a relative small value means certain state estimates [26, 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, representing 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 (12).

3.2. Method 2 – noise in control inputs

Assume that the system (15)–(18) is considered a semi-batch system in that the species B is removed and that the species A is refilled when a certain low level of A (represented by the partial pressure P_A) is reached (as described by (28)). Consider also that there is some uncertainty related to the control inputs, and that the uncertainty in the inputs can be expressed as a relative uncertainty. The model for the estimator becomes

$$\dot{\hat{x}}_1 = -2\hat{r} + u_1(1 + v_{u_1}) \tag{30}$$
$$\dot{\hat{x}}_2 = \hat{r} + u_2(1 + v_{u_2}) \tag{31}$$
$$\dot{\hat{x}}_3 = v_{k_r} \tag{32}$$

where the state vector $\hat{x}$ is given by (19) and the reaction rate $\hat{r}$ is given by

$$\hat{r} = \hat{k}_r \hat{x}_1^2 = \hat{x}_3 \hat{x}_1^2 \tag{33}$$

The noise vector is given by

$$v = \begin{bmatrix} v_{u_1} & v_{u_2} & v_{k_r} \end{bmatrix}^T \tag{34}$$

where $v_{u_1} \sim N(0, \bar{v}_{u_1}^2)$, $v_2 \sim N(0, \bar{v}_{u_2}^2)$ and $v_{k_r} \sim N(0, \bar{v}_{k_r}^2)$. The parameters are given by (27). The Jacobian G becomes

---

[4] Note that the in the actual implementation, using a discrete formulation, the control input is scaled such that the refilling/removing is done in one sample. Note also that the measured state is used for control purposes, since it is the estimator that is under investigation, and not the output feedback system.

$$G = \begin{bmatrix} u_1 & 0 & 0 \\ 0 & u_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (35)$$

Assuming Q is described by

$$Q = \begin{bmatrix} \bar{v}_{u_1}^2 & 0 & 0 \\ 0 & \bar{v}_{u_2}^2 & 0 \\ 0 & 0 & \bar{v}_{k_r}^2 \end{bmatrix} \quad (36)$$

the term GQG^T in (12) becomes

$$GQG^T = \begin{bmatrix} u_1^2 \bar{v}_{u_1}^2 & 0 & 0 \\ 0 & u_2^2 \bar{v}_{u_2}^2 & 0 \\ 0 & 0 & \bar{v}_{k_r}^2 \end{bmatrix} \quad (37)$$

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

3.3. Method 3 – noise in auxiliary variables

Consider the system (15)–(18) and that process knowledge suggest that the noise enters the system through the reaction rate r, instead of directly on the system states. That is

$$\dot{x}_1 = -2r \quad (38)$$
$$\dot{x}_2 = r \quad (39)$$

where the state vector x is given by (18) and the reaction rate r is given by

$$r = k_r x_1^2 + v_r \quad (40)$$

The model for the estimator becomes

$$\dot{\hat{x}}_1 = -2\hat{r} \quad (41)$$
$$\dot{\hat{x}}_2 = \hat{r} \quad (42)$$
$$\dot{\hat{x}}_3 = v_{k_r} \quad (43)$$

where the state vector $\hat{x}$ is given by (19) and the reaction rate $\hat{r}$ is given by

$$\hat{r} = \hat{k}_r \hat{x}_1^2 + v_r = \hat{x}_3 \hat{x}_1^2 + v_r \quad (44)$$

The noise vector is given by

$$v = [\, v_r \quad v_{k_r} \,]^T \quad (45)$$

where $v_r \sim N(0, \bar{v}_r^2)$ and $v_{k_r} \sim N(0, \bar{v}_{k_r}^2)$. The parameters are given by (27). The system becomes

$$\dot{\hat{x}}_1 = -2\hat{k}_r \hat{x}_1^2 - 2v_r \quad (46)$$
$$\dot{\hat{x}}_2 = \hat{k}_r \hat{x}_1^2 + v_r \quad (47)$$
$$\dot{\hat{x}}_3 = v_{k_r} \quad (48)$$

and the Jacobian G

$$G = \begin{bmatrix} -2 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (49)$$

Assuming Q is described by

$$Q = \begin{bmatrix} \bar{v}_r^2 & 0 \\ 0 & \bar{v}_{k_r}^2 \end{bmatrix} \quad (50)$$

the term GQG^T in (12) becomes

$$GQG^T = \begin{bmatrix} 4\bar{v}_r^2 & -2\bar{v}_r^2 & 0 \\ -2\bar{v}_r^2 & \bar{v}_r^2 & 0 \\ 0 & 0 & \bar{v}_{k_r}^2 \end{bmatrix} \quad (51)$$

That is, by applying noise on the auxiliary variable $\hat{r}$, correlation is naturally introduced in the covariance calculation (P). This could also be seen as an alternative to off-diagonal tuning on Q in (29), in that the correlation enters the system naturally and correctly scaled in the off-diagonals.

4. Simulation studies

In the following sections results with some of the noise methods described above are presented. Note that modeling error may be introduced via θ_k and to some extent through the noise inputs v_k and w_k. Further, the noise inputs are generated using the randn( ) function in Matlab. Note also that in all the simulations a discrete representation based on Runge–Kutta 4th order integration scheme is used.

The first case is from Hasseltine and Rawlings [9], where it was used to indicate the convergence performance of the Moving Horizon Estimation (MHE) approach and at the same time show the limitations using the EKF. The key challenge here is that it has a multi modal probability density function, where some solutions are physically valid, and some not. Given a bad initial guess, an algorithm based on the Kalman filter approach may or may not converge to the true state.

The second case shows results from simulations done on an industrial simulator of the Hall–Heroult process.

The third case is a reactor case, inspired by the problem described by Rawlings et al. [21], investigating one of Kodak's semi-batch speciality chemicals for photographic film production. The case appears in Rao [20], and is used to compare the performance of the EKF and MHE.

All simulations were done in Matlab Release 14. The estimator used in this work is the fully augmented UKF with reformulation of the correction steps and the use of constraints as presented in Kolås et al. [16], hereafter denoted CUKF. Further, when parameter estimation is required, the CUKF used is an extended version allowing parameter estimation (Joint CUKF[5]). Note that the UKF requires a discrete representation of the system.

To cope with singular/negative definite covariance matrix with respect to matrix square root calculations, we have introduced the term δ_i such that

$$\begin{aligned} P_{x_k} &= P_{x_k}^- - \delta_k K P_{y_k y_k}^- K^T \\ \Omega_i &= [0, 0.1, 0.2, 0.3, \ldots, 1] \\ \delta_k &= \max(\Omega_i)|_{P_{x_k} > 0} \end{aligned} \quad (52)$$

If $P_{x_k}^- > 0$ then so is P_{x_k}. Note also that scaling is not an issue in any of the following simulations.

4.1. Case '2 state CSTR'

4.1.1. Case description
In the following, the systems as described by (15)–(28) are investigated. Further, the parameters for these systems, if not otherwise noted, is

---

[5] The augmented CUKF/UKF is in relevant literature referred to as augmentation of the noise matrices in the sigma points, hence the use of Joint to state that parameters are augmented into the state vector.

$\Delta t = t_{k+1} - t_k = 0.1$

$$P_0 = \begin{bmatrix} 6^2 & 0 & 0 \\ 0 & 6^2 & 0 \\ 0 & 0 & 0.015^2 \end{bmatrix}$$

$k_r = 0.16$

$$x_0 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \tag{53}$$

$$\hat{x}_0 = \begin{bmatrix} 0.1 \\ 4.5 \\ 0.9k_r \end{bmatrix}$$

Note that the initial guess $(\hat{x}_0)$ for the estimator is poor.

First investigated is a case where the reaction parameter k_r is constant. It has a wrong initial value, however. Then follows an investigation of a case with step changes in the reaction parameter k_r. The step changes in k_r are invoked as

$$\begin{array}{ccccc} t & 0 & 46 & 200 & 400 \\ k_r & 0.16 & 0.12 & 0.17 & 0.16 \end{array} \tag{54}$$

The following constraints are applied to the CUKF sigma points

$Lowerbounds: \ [0, 0, 0.1]^T$

$Upperbounds: \ [\infty, \infty, 0.18]^T$ \hfill (55)

Note that the actual CUKF applied is a Joint CUKF, since the state vector is augmented in order to perform parameter estimation of k_r. Note also that for all the simulations the simulator is based on a discrete representation of (20)-(21).

4.1.2. Simulation results with constant k_r

Method 1.

The system equations and the noise is modeled as given in the description of Method 1, the estimator constraints as (55) and the estimator tuning as

$$Q_k = diag[\ \bar{v}_{k,1}^2 \quad \bar{v}_{k,2}^2 \quad \bar{v}_{k,k_r}^2 \ ]$$
$$= diag[\ (10^{-9})^2 \quad (10^{-9})^2 \quad (10^{-4})^2 \ ] \tag{56}$$
$$R_k = \bar{w}_k^2 = 0.002^2$$

In the case when the true process, i.e. the simulator, experience no noise, we get the results as shown in Fig. 1.

As Fig. 1 shows, the bad initial guess is handled very well, and the state estimate and output (measurement) estimate is generally acceptable. However the parameter estimate during the transient period at approximately 'Time 15' is generally bad due to the control input excitation. In Fig. 2 the time trace of the covariance is shown, and as expected the control input excitation is not reflected in the covariance.

Methods 1 and 2 combined. By combining the concept described by Methods 1 and 2, and by that introduce uncertainty also in the control inputs, the system equations for the estimator becomes

$$\dot{\hat{x}}_1 = -2\hat{r} + u_1(1 + v_{u_1}) + v_1 \tag{57}$$
$$\dot{\hat{x}}_2 = \hat{r} + u_2(1 + v_{u_2}) + v_2 \tag{58}$$
$$\dot{\hat{x}}_3 = v_{k_r} \tag{59}$$

where the state vector $\hat{x}$ is given by (19) and the reaction rate $\hat{r}$ is given by

$$\hat{r} = \hat{k}_r \hat{x}_1^2 = \hat{x}_3 \hat{x}_1^2 \tag{60}$$

The noise is given by $v_1 \sim N(0, \bar{v}_1^2)$, $v_2 \sim N(0, \bar{v}_2^2)$, $v_{k_r} \sim N(0, \bar{v}_{k_r}^2)$, $v_{u_1} \sim N(0, \bar{v}_{u_1}^2)$, $v_{u_2} \sim N(0, \bar{v}_{u_2}^2)$ and the parameters as by (27). Further, by assuming that Q is described by

$$Q = \begin{bmatrix} \bar{v}_1^2 & 0 & 0 & 0 & 0 \\ 0 & \bar{v}_2^2 & 0 & 0 & 0 \\ 0 & 0 & \bar{v}_{k_r}^2 & 0 & 0 \\ 0 & 0 & 0 & \bar{v}_{u_1}^2 & 0 \\ 0 & 0 & 0 & 0 & \bar{v}_{u_1}^2 \end{bmatrix} \tag{61}$$
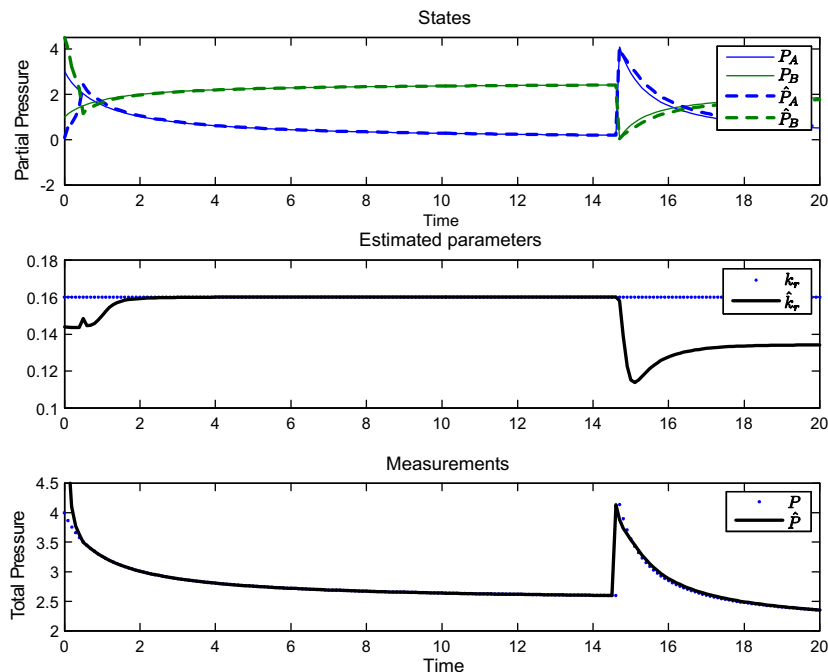
the term GQG^T in (12) becomes

Fig. 1. 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 period caused by the control input excitation.
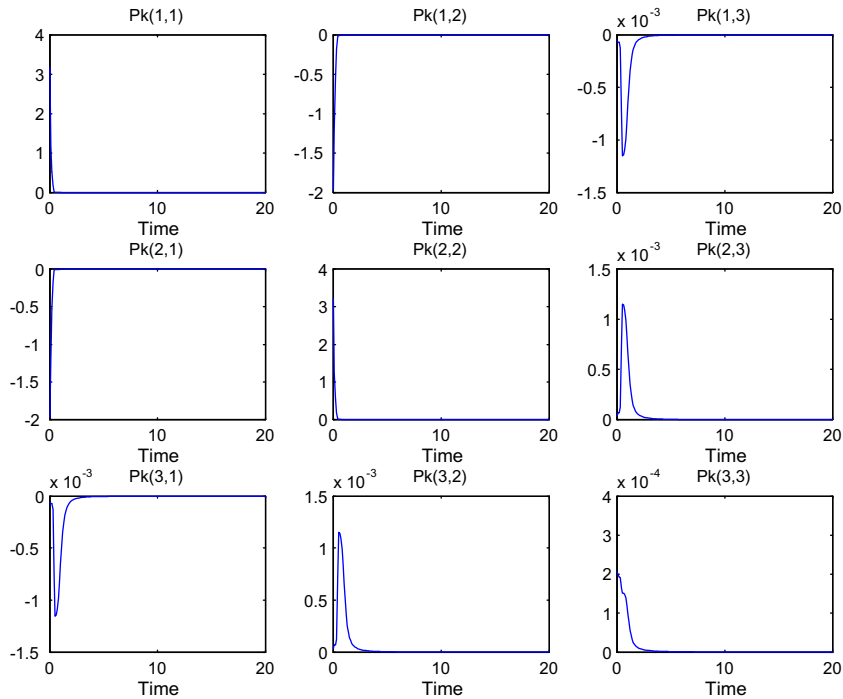
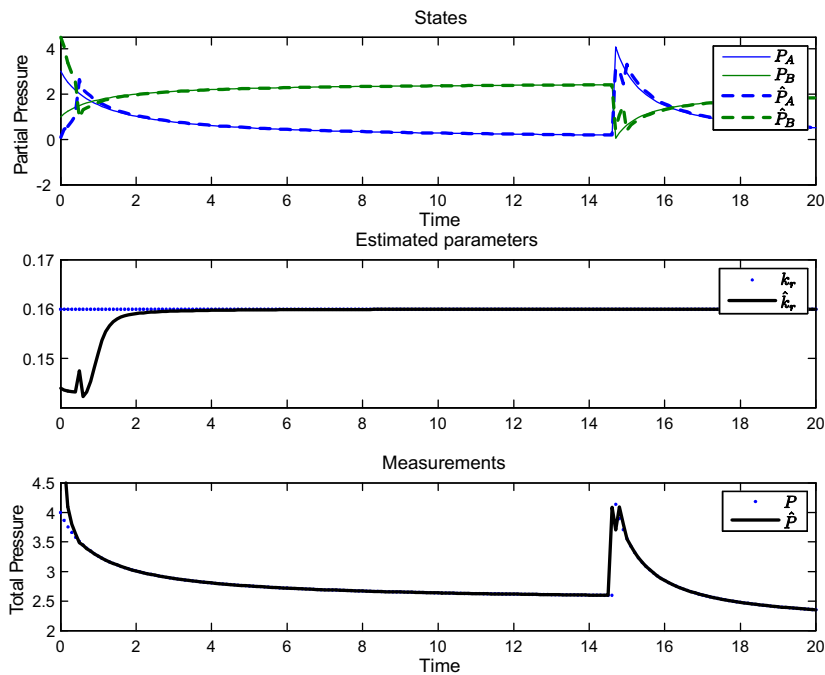**Fig. 2.** The time trace of the discrete covariance matrix P_{x_k}.

Fig. 3. 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 period caused by the control input excitation.

$$GQG^T = \begin{bmatrix} \bar{v}_1^2 + u_1^2 \bar{v}_{u_1}^2 & 0 & 0 \\ 0 & \bar{v}_2^2 + u_2^2 \bar{v}_{u_2}^2 & 0 \\ 0 & 0 & \bar{v}_{k_r}^2 \end{bmatrix} \quad (62)$$

and as seen by (62), the control inputs will be present in the covariance equation.

With the estimator constraints as by (55), and the estimator tuning as

$$Q_k = diag[\, \bar{v}_{k,1}^2 \quad \bar{v}_{k,2}^2 \quad \bar{v}_{k,k_r}^2 \quad \bar{v}_{k,u_1}^2 \quad \bar{v}_{k,u_2}^2 \,]$$
$$= diag[\,(10^{-9})^2 \quad (10^{-9})^2 \quad (10^{-4})^2 \quad 1^2 \quad 1^2 \,] \quad (63)$$
$$R_k = \bar{w}_k^2 = 0.002^2$$

the results for the case when the simulator experience no noise, is as shown in Fig. 3.

As Fig. 3 shows, the estimate of the parameter k_r does not suffer from the control input excitation. However the 'cost' is some loss
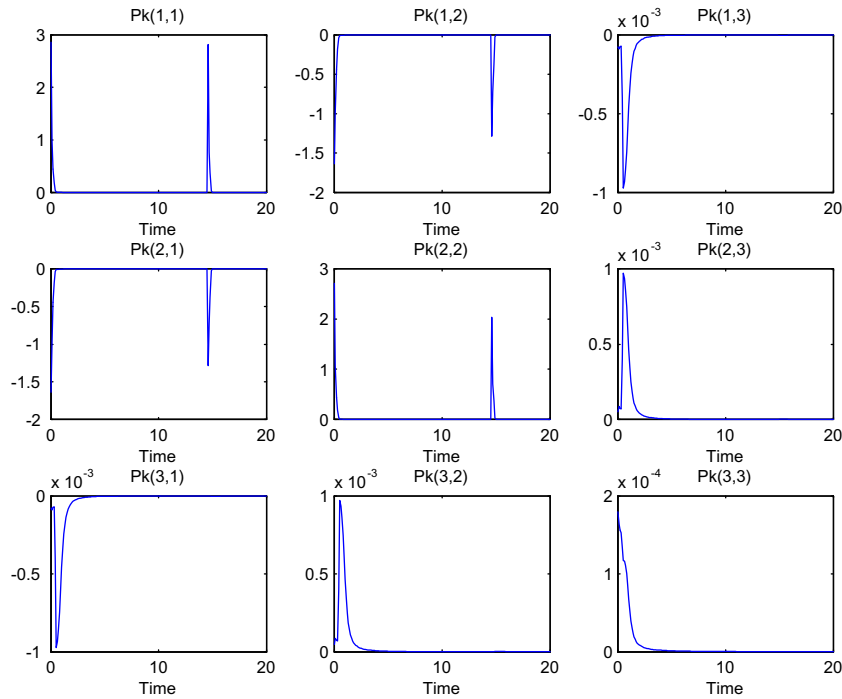
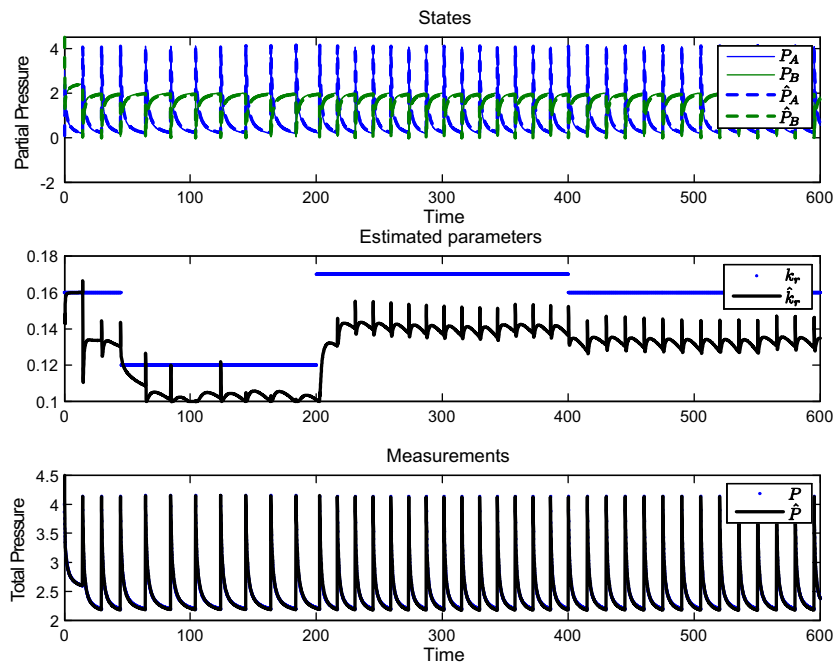**Fig. 4.** The time trace of the discrete covariance matrix P_{x_k}.

Fig. 5. The figure shows the performance of the estimator using the noise formulation given in Method 1. The errors in the states are relative small, but the performance is bad with respect to how well the estimator suppresses the control input excitation and the large offset in the estimation of k_r.

off accuracy regarding the state estimates in the first samples after the control input excitation. In Fig. 4 the time trace of the covariance is shown, and as expected the control input excitation at approximately 'Time 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 Methods 1 and 2.

4.1.3. Simulation results with time-varying k_r

Method 1. Again we consider the system given by (22)–(26). That is, the traditional noise model (Method 1) and with the tuning of the estimator as in (55)-(56). The constraints regarding the sigma points χ_{k-1}^{x-} and χ_k^x is as in the case previously described. Further the parameter k_r changes according to (54). The result of the simulation is given in Fig. 5.

As Fig. 5 shows, the changes in the true reaction parameter k_r is to some extent reflected, but with a bias and some noise from the
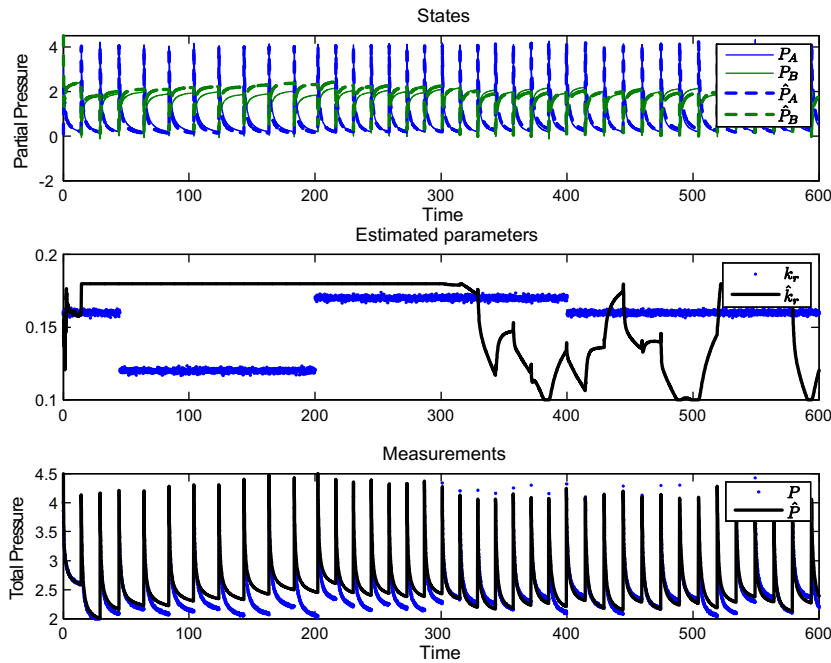
Fig. 6. The figure shows the performance of the estimator using the noise formulation given in Method 1. The errors in the states are large, and the performance is bad with respect to the accuracy of the estimated k_r as well as poor suppression of the control input excitations.
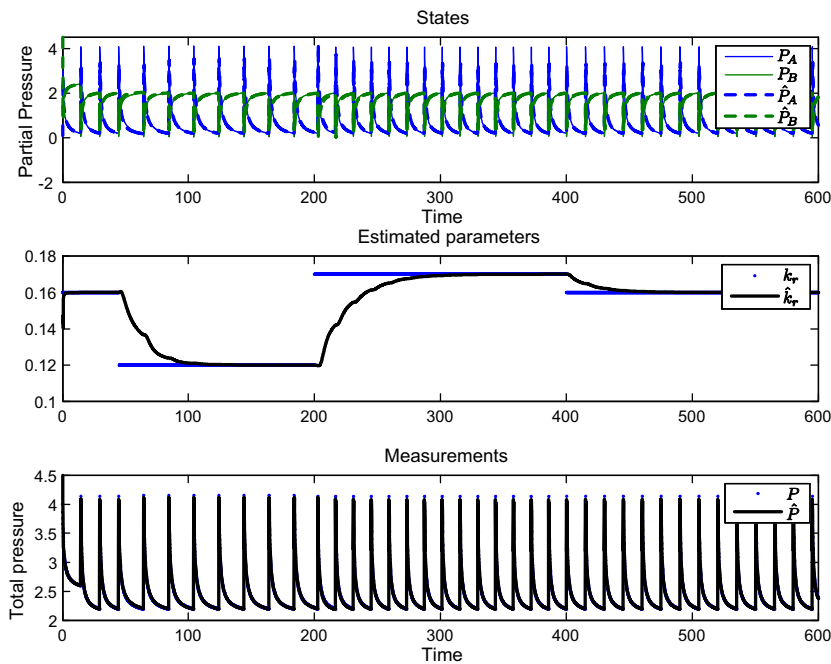

Fig. 7. The figure shows the performance of the estimator using the noise formulation combining Methods 1 and 2. The errors in the state estimates are small, and the performance good with respect to how well the estimator suppress the control input excitation and how well the estimation of k_r converge to the true state.

input excitation. However, Fig. 5 reflects the case when the true process experiences no noise. By letting the true system experiences noise as

$$v \sim N\left(0^T, \; [\; \bar{v}_1^2 \quad \bar{v}_2^2 \quad \bar{v}_{k_r}^2 \quad \bar{v}_{u_1}^2 \quad \bar{v}_{u_2}^2 \;]^T\right)$$

$$= N\left(0^T, [0.001^2 \quad 0.001^2 \quad 0.001^2 \quad 0.1^2 \quad 0.1^2]^T \Delta t^{-1}\right) \tag{64}$$

$$w \sim N(0, \bar{w}^2) = N(0, 0.01^2 \Delta t) \tag{65}$$

and the estimator constraints as by (55) and estimator tuning as given by (56), the results are as shown in Fig. 6.

As Fig. 6 shows, the estimator performs bad given this tuning and noise sequence.

Methods 1 and 2 combined. By combining Methods 1 and 2, i.e. consider the system with additive noise and assume relative noise in the control inputs to let the covariance reflect the uncertainty in the state estimates when at the control input excitation. The esti-
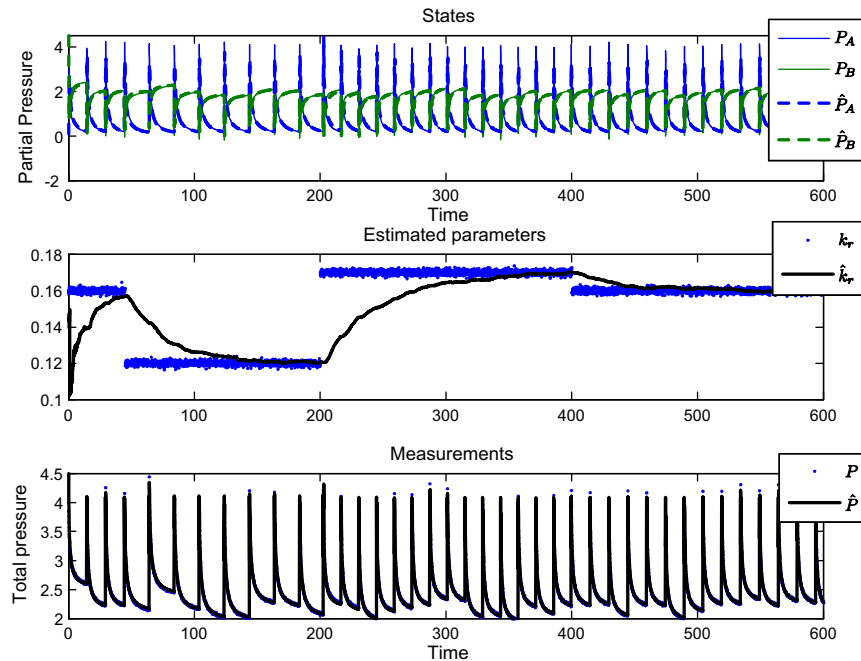
Fig. 8. The figure shows the performance of the estimator using the noise formulation combining Methods 1 and 2. The errors in the state estimates are relatively small, and the performance good with respect to how well the estimator suppress the control input excitation. Also, the estimator manages to estimate k_r reasonably well.
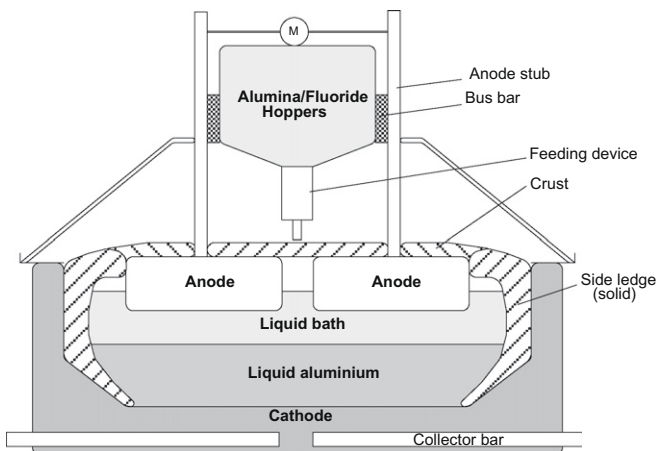

Fig. 9. The figure illustrates a cell for producing liquid aluminium.

mator tuning is as given by (55) and (63). The results are shown in Fig. 7.

As Fig. 7 shows the performance in the noise free case is acceptable. By letting the true system experience noise as given by (64) and (65), and the estimator tuning as by (55) and (63), we get the results as given by Fig. 8.

As Fig. 8 shows, given this particular noise sequence and particular noise levels, the estimator performs well with respect to the state estimates. The estimate of the parameter k_r is reasonably well, converging slowly towards the true state. Also, the control input excitations are no longer disturbing the estimate of k_r.

4.2. Case 'Hall–Heroult Process'

4.2.1. Case description

The[6] Hall–Heroult process is the dominating process for producing aluminum today [8]. The fundamentals of the process are to dis-

solve Al_2O_3 in molten cryolite (also known as electrolyte or bath), and electrically reduce complex aluminum containing ions to pure aluminum. The overall electrochemical reaction in the electrolyte is

$$2Al_2O_3 + 3C \rightarrow 4Al + 3CO_2$$

where carbon is fed to the reaction as consumable anodes. By the use of various additives, in particular AlF_3, the operating temperature of the electrolyte can be lowered from 1010 °C to approximately 960 °C. Both decreased temperature and increased excess AlF_3 is believed to be beneficial for the amount of metal produced (current efficiency) and the energy consumption. As molten cryolite is very corrosive, the only component of an acceptable cost presently capable of coexisting with it over time is frozen cryolite. It is therefore necessary to maintain a layer of frozen cryolite (side ledge) to prevent the carbon walls from eroding. In order to maintain the side ledge there has to be a substantial heat loss through the side ledge and the carbon walls of the cell. The cell voltage applied is typically between 4.2 and 4.5 V, and the electric current through the cell is typically 150–350 kA. A sketch of a cell is shown in Fig. 9. In a modern plant of today 100–300 cells are placed and connected in series. There are three control inputs to the process, anode beam adjustments (controlling energy input), addition of AlF_3 and addition of Al_2O_3, and three controlled variables, electrolyte temperature,[7] concentration (or mass) of AlF_3 and concentration of Al_2O_3. A cell is regularly excited since liquid aluminium is tapped and some of the anode blocks are changed on a daily basis. This induces severe disturbances in the energy balance, and it implies that the operating conditions will vary significantly and hence provoke nonlinear cell effects. 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 [5,4,7].

In this work we investigate, using an industrial simulator, the Hall–Heroult process in closed loop control by an advanced process

---

[6] The case description is mainly from Foss and Schei [5].
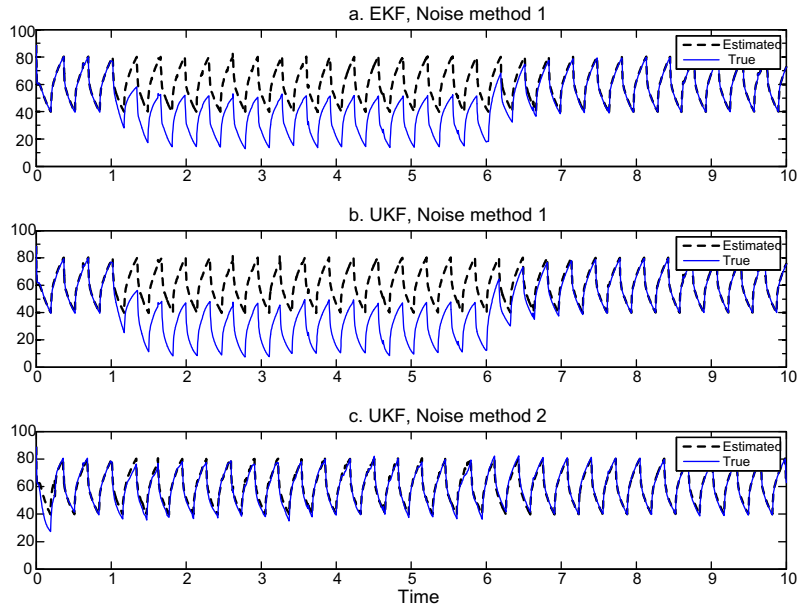
[7] Also known as bath temperature.

Fig. 10. The figure indicates the effect of using the noise concepts as described by Method 1 (subplot a and b) and the combination of Methods 1 and 2 (subplot c) on a simulated industrial process which experience input uncertainty.

control application as used by Hydro Aluminium AS. The model used is an adoption of the models found in Saksvikronning et al. [22] and Gran [7]. The estimators used in the simulations are the EKF (Gran [7]) and the fully augmented UKF [29].

As already mentioned, a particular challenge with the Hall–Heroult process is that the control inputs may experience severe disturbances. In the simulation a reduction of 5% in one of the control inputs is introduced in the simulator. The drop in the control input is not known to the control application. The purpose of the simulation is to study the noise disturbance models described by Methods 1 and 2 on this process.

4.2.2. Simulation results

Fig. 10a–c shows the results of the simulation. The 5% drop in the control input is initiated at time stamp 1 and last until time stamp 6. After time stamp 6 the control input is considered back to normal. Note that the data is anonymized due to confidentiality issues.

Fig. 10 indicates that the drop in the true control input at time step 1 is not 'discovered' by the estimator, either using EKF or UKF applying the noise modeling concept described by Method 1. However applying the noise modeling concept by combining Methods 1 and 2 (i.e. input uncertainty), the UKF estimator 'discover' the drop, which is indicated in Fig. 10 subplot c, by no drop in the true state (remember the system is in closed loop control). As indicated by Fig. 10 subplot a and b, the drop in the control input is not discovered by the estimators using only the traditional additive noise formulation. Note that a simulation regarding how the EKF combined with noise combining Methods 1 and 2 in this case was not performed, but is expected to behave similar to the UKF as shown in Fig. 10, subplot c.

4.3. Case '4 State CSTR'

4.3.1. Case description

Consider a stirred tank reactor where the following liquid phase exothermic reaction occurs

$$A + 2B \rightarrow C. \tag{66}$$

This case is found in Rao [20],[8] which is inspired by the problem described by Rawlings et al. [21]. The state estimation problem is to estimate precisely the concentration of A in the reactor. Because over-addition of B leads to product degradation, precise concentration estimates of A as a function of time are necessary to complete the reaction without over-addition of B. Only temperature measurements corrupted with sensor noise are available. Further, the exact reaction kinetics are supposed unknown with the exception of the heat of reaction ΔH_r [24]. Under standard assumptions, such as negligible potential and kinetic energy effects, constant density, uniformly homogeneous mixture, and no phase transition, the reactor is simulated using the following model [28]:

$$\dot{V} = F \tag{67}$$

$$\dot{A} = -k_0 e^{(-\frac{ER}{T})} AB^2 - \frac{F}{V} A \tag{68}$$

$$\dot{B} = -2k_0 e^{(-\frac{ER}{T})} AB^2 + \frac{F}{V} \left(C_{B_f} - B \right) \tag{69}$$

$$\dot{T} = -\frac{\Delta H_r}{\rho C_p} k_0 e^{(-\frac{ER}{T})} AB^2 + \frac{F}{V} (T_f - T) + \frac{UA}{\rho C_p V} (T_c - T) \tag{70}$$

The model parameters are listed in Table 1.

The flowrate profile (F) used, is a rescaled version of the one used in the operation of the industrial reactor described by Rawlings et al. [21]. To account for imperfect cooling and modeling inaccuracies the cooling water temperature (T_c) is assumed fluctuating. The control inputs[9] $F(u_1)$ and the measured disturbance $T_c(u_2)$ is as given in Fig. 11.

By introducing the notation for the states x and control inputs u as

---

[8] Note that the sampling time ($h = 0.1$ [min]) used in this work differ from the one ($h = 0.5$ [min]) used in Rao [20], because the simulator used in this work experience instabilities if $h > 0.25$ [min].

[9] Note that the control input u_1 is a scaled version of the one given in Rawlings et al. [21]. The 'measured' temperature of the coolingwater $T_c \sim N(\bar{T}_c, v_{T_c}^2) = N(300, 5)$ is generated based on the original sampling rate as u_1 and then scaled as in Rao [20].

Table 1
The table explains the parameters used in the 4-state CSTR case.

Parameter	Value	Units
k_0	$9 \cdot 10^{11}$	$\frac{L^2}{mol^2 \cdot min}$
ER	6000	K
ρ	1000	$\frac{g}{L}$
C_p	0.239	$\frac{J}{g \cdot K}$
UA	$2 \cdot 10^5$	$\frac{J}{min \cdot K}$
T_f	300	K
T_c	300	K
C_{B_f}	2.2	$\frac{mol}{L}$
ΔH_r	$-5 \cdot 10^4$	$\frac{J}{mol}$
F	–	$\frac{L}{min}$
r	–	$\frac{mol}{min \cdot L}$
Q_r	–	$\frac{K}{min}$
$V(0)$	100	L
$A(0)$	0.5	$\frac{mol}{L}$
$B(0)$	0	$\frac{mol}{L}$
$T(0)$	300	K
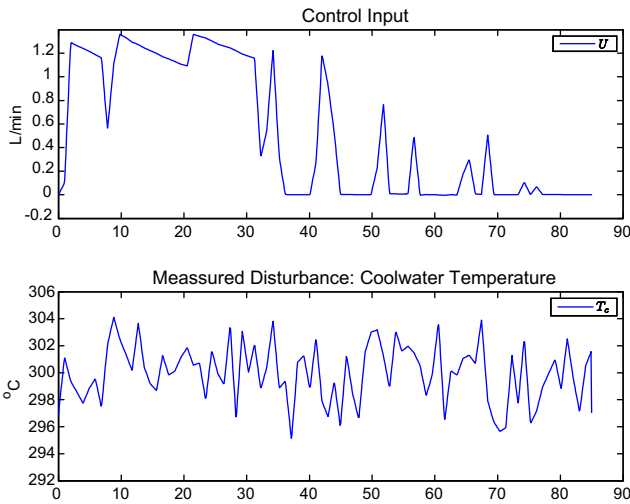
Fig. 11. The control input u_1 and the measured cooling water temperature T_c.

$$x = [x_1 \quad x_2 \quad x_3 \quad x_4]^T = [V \quad A \quad B \quad T]^T \tag{71}$$

$$u = [u_1 \quad u_2]^T = [F \quad T_c]^T, \tag{72}$$

the model can be formulated as

$$\dot{x}_1 = u_1 \tag{73}$$

$$\dot{x}_2 = -k_0 e^{\left(-\frac{ER}{x_4}\right)} x_2 x_3^2 - \frac{x_2}{x_1} u_1 \tag{74}$$

$$\dot{x}_3 = -2k_0 e^{\left(-\frac{ER}{x_4}\right)} x_2 x_3^2 + \frac{\left(C_{B_f} - x_3\right)}{x_1} u_1 \tag{75}$$

$$\dot{x}_4 = -\frac{\Delta H_r}{\rho C_p} k_0 e^{\left(-\frac{ER}{x_4}\right)} x_2 x_3^2 + \frac{UA}{\rho C_p} \frac{(u_2 - x_4)}{x_1} + \frac{(T_f - x_4)}{x_1} u_1 \tag{76}$$

Defining the reaction rate r as

$$r = k_0 e^{\left(\frac{ER}{x_4}\right)} x_2 x_3^2 \tag{77}$$

gives the model

$$\dot{x}_1 = u_1 \tag{78}$$

$$\dot{x}_2 = -r - \frac{x_2}{x_1} u_1 \tag{79}$$

$$\dot{x}_3 = -2r + \frac{\left(C_{B_f} - x_3\right)}{x_1} u_1 \tag{80}$$

$$\dot{x}_4 = -\frac{\Delta H_r}{\rho C_p} r + \frac{UA}{\rho C_p} \frac{(u_2 - x_4)}{x_1} + \frac{(T_f - x_4)}{x_1} u_1. \tag{81}$$

Further it is assumed that the reaction rate r is unknown and we will investigate two different estimation approaches addressing this issue. Both approaches assumes that the reaction rate r could be modeled as colored noise. First we assume that only a simplified model of the reaction rate is available (as in Rao [20]), and that the reaction rate r is derived from the estimated reaction energy Q_r. The estimator model takes the form

$$\dot{\hat{x}}_1 = u_1 \tag{82}$$

$$\dot{\hat{x}}_2 = -\hat{r} - \frac{\hat{x}_2}{\hat{x}_1} u_1 \tag{83}$$

$$\dot{\hat{x}}_3 = -2\hat{r} + \frac{\left(C_{B_f} - \hat{x}_3\right)}{\hat{x}_1} u_1 \tag{84}$$

$$\dot{\hat{x}}_4 = -\frac{\Delta H_r}{\rho C_p} \hat{r} + \frac{UA}{\rho C_p} \frac{(u_2 - \hat{x}_4)}{\hat{x}_1} + \frac{(T_f - \hat{x}_4)}{\hat{x}_1} u_1 \tag{85}$$

$$\dot{\hat{x}}_5 = v_{x_5} \tag{86}$$

where

$$\hat{x}_5 = \hat{Q}_r = -\frac{\Delta H_r}{\rho C_p} \hat{r} \tag{87}$$

and control inputs u as (72) and the states $\hat{x}$ as

$$\hat{x} = [\hat{x}_1 \quad \hat{x}_2 \quad \hat{x}_3 \quad \hat{x}_4 \quad \hat{x}_5]^T = [\hat{V} \quad \hat{A} \quad \hat{B} \quad \hat{T} \quad \hat{Q}_r]^T \tag{88}$$

Next, we assume that knowledge on how the states enter the reaction rate (77) is available, but where the true value of the reaction rate constant k_0 is uncertain and must be estimated. In this case $\hat{x}_5$ in (86) takes the form

$$\hat{x}_5 = \hat{k}_0 \tag{89}$$

and the reaction rate by

$$\hat{r} = \hat{k}_0 e^{\left(-\frac{ER}{\hat{x}_4}\right)} \hat{x}_2 \hat{x}_3^2 \tag{90}$$

and the control inputs u as (72) and the states $\hat{x}$ as

$$\hat{x} = [\hat{x}_1 \quad \hat{x}_2 \quad \hat{x}_3 \quad \hat{x}_4 \quad \hat{x}_5]^T = [\hat{V} \quad \hat{A} \quad \hat{B} \quad \hat{T} \quad \hat{k}_0]^T \tag{91}$$

Note that for all the simulations to come, it is assumed that the true system experiences noise in x_1 and that the noise in the other states enters via noise in the reaction rate r, i.e.

$$r = k_0 e^{\left(-\frac{ER}{x_4}\right)} x_2 x_3^2 + v_r \tag{92}$$

where

$$v_r \sim N(0, \bar{v}_r^2) = N\left(0, 10^{-6}/\Delta t\right) \tag{93}$$

$$v_{x_1} \sim N(0, \bar{v}_{x_1}^2) = N(0, 10^{-10}/\Delta t) \tag{94}$$

Further the true measurement noise is assumed to be

$$w \sim N(0, \bar{w}_T^2) = N(0, 0.1^2 \Delta t) \tag{95}$$

4.3.2. Method 1, simplified model

We investigate the behavior of the estimator when erroneously assuming the standard approach that all the states experience additive noise, and no noise in the auxiliary parameters. The state vector is as given by (88) and the estimator tuning as

$$\hat{x}_0 = [\,100 \quad 0.5 \quad 0 \quad 300 \quad 0\,]^T \tag{96}$$

$$P_0 = diag[\,10^{-5} \quad 10^{-5} \quad 10^{-5} \quad 10^{-5} \quad 10^{-1}\,] \tag{97}$$

$$Q_k = diag[\,\bar{v}_{x_1}^2 \quad \bar{v}_{x_2}^2 \quad \bar{v}_{x_3}^2 \quad \bar{v}_{x_4}^2 \quad \bar{v}_{x_5}^2\,]$$

$$= diag[\,(10^{-5})^2 \quad (10^{-5})^2 \quad (7\cdot10^{-2})^2 \quad (10^{-5})^2 \quad 1^2\,] \tag{98}$$

$$R_k = \bar{w}_k^2 = 0.075^2 \tag{99}$$

Sigma points lower bounds : $[\,0, \quad 0, \quad 0, \quad 0, \quad 0\,]^T$
Sigma points upper bounds : $[\,\infty, \quad \infty, \quad 0.011, \quad \infty, \quad 3\,]^T$ (100)

Using the standard approach with additive noise on the states, the results are as given in Fig. 12.

As is seen by Fig. 12, the estimation of the concentration of A is close to the true one, but the estimated reaction rate and the estimated concentration of B suffer from noise. In this case we have used v_{x_3} to tune the behavior of the estimated concentration of B.

4.3.3. Methods 1 and 3 combined, simplified model

Applying 'process knowledge' we may assume that the reaction rate r is uncertain and experience noise as

$$\hat{Q}_r = -\frac{\Delta H_r}{\rho C_p}(\hat{r} + v_r) \tag{101}$$

$v_r \sim N(0, \bar{v}_r^2)$ and the reaction rate becomes

$$\hat{r} = -\hat{Q}_r\left(\frac{\Delta H_r}{\rho C_p}\right)^{-1} - v_r = -c_1\hat{x}_5 - v_r \tag{102}$$

$$c_1 = \left(\frac{\Delta H_r}{\rho C_p}\right)^{-1} \tag{103}$$

The model for the estimator becomes

$$\dot{\hat{x}}_1 = u_1 + v_{x_1} \tag{104}$$

$$\dot{\hat{x}}_2 = -\hat{r} - \frac{\hat{x}_2}{\hat{x}_1}u_1 \tag{105}$$

$$\dot{\hat{x}}_3 = -2\hat{r} + \frac{\left(C_{B_f} - \hat{x}_3\right)}{\hat{x}_1}u_1 \tag{106}$$

$$\dot{\hat{x}}_4 = -c_1^{-1}\hat{r} + \frac{UA}{\rho C_p}\frac{(u_2 - \hat{x}_4)}{\hat{x}_1} + \frac{(T_f - \hat{x}_4)}{\hat{x}_1}u_1 \tag{107}$$

$$\dot{\hat{x}}_5 = v_{x_5} \tag{108}$$

The noise vector v is defined as

$$v = [\,v_{x_1} \quad v_r \quad v_{x_5}\,]^T \tag{109}$$

where $v_{x_1} \sim N(0, \bar{v}_{x_1}^2), v_r \sim N(0, \bar{v}_r^2)$ and $v_{x_5} \sim N(0, \bar{v}_{x_5}^2)$. By this H becomes

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & c_1^{-1} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{110}$$

and GQG^T

$$GQG^T = \begin{bmatrix} \bar{v}_{x_1}^2 & 0 & 0 & 0 & 0 \\ 0 & \bar{v}_r^2 & 2\bar{v}_r^2 & c_1^{-1}\bar{v}_r^2 & 0 \\ 0 & 2\bar{v}_r^2 & 4\bar{v}_r^2 & 2c_1^{-1}\bar{v}_r^2 & 0 \\ 0 & c_1^{-1}\bar{v}_r^2 & 2c_1^{-1}\bar{v}_r^2 & c_1^{-1}\bar{v}_r^2 & 0 \\ 0 & 0 & 0 & 0 & \bar{v}_{x_5}^2 \end{bmatrix} \tag{111}$$

when

$$Q = diag[\,\bar{v}_{x_1}^2 \quad \bar{v}_r^2 \quad \bar{v}_{x_5}^2\,] \tag{112}$$

For the estimator the state vector is as given by (88) and the other initial values are

$$\hat{x}_0 = [\,100 \quad 0.5 \quad 0 \quad 300 \quad 0\,]^T \tag{113}$$

Sigma points lower bounds : $[\,0, \quad 0, \quad 0, \quad 0, \quad 0\,]^T$
Sigma points upper bounds : $[\,\infty, \quad \infty, \quad 0.011, \quad \infty, \quad 3\,]^T$ (114)

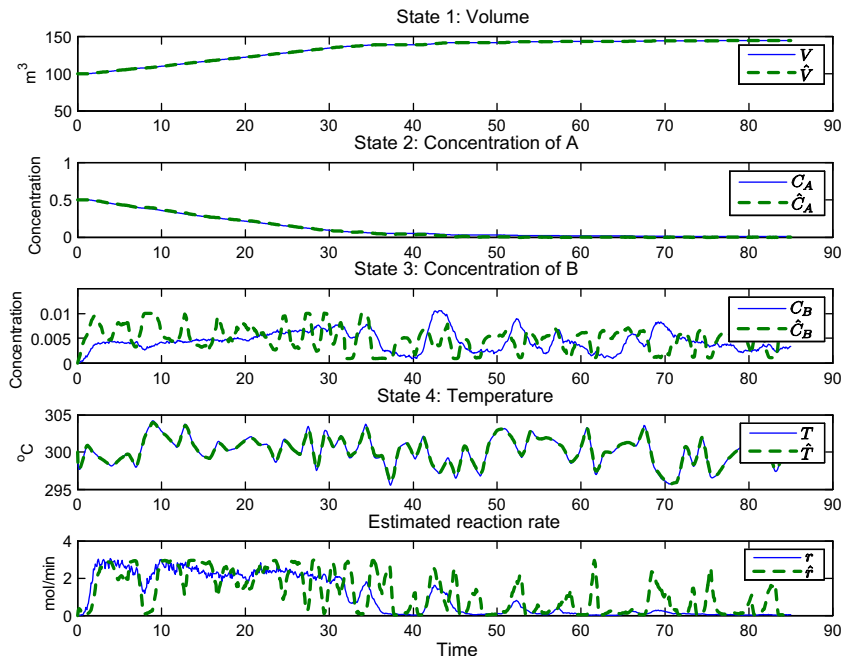$$P_0 = diag[\,10^{-5} \quad 10^{-5} \quad 10^{-5} \quad 10^{-5} \quad 10^{-1}\,] \tag{115}$$

Fig. 12. The results of the simulation assumed simplified model and additive noise.

$$Q_k = diag[\, \bar{v}_{k,x_1}^2 \quad \bar{v}_{k,r}^2 \quad \bar{v}_{k,x_5}^2 \,] = diag\left[\, \left(10^{-5}\right)^2 \quad \left(10^{-3}\right)^2 \quad \left(1.1 \cdot 10^{11}\right)^2 \,\right] \tag{116}$$

$$R_k = \bar{w}_k^2 = 0.075^2 \tag{117}$$

The simulation results are as given in Fig. 13.

As is seen by Fig. 13, the estimation of the concentration of A is close to the true one (marginally better than for the additive noise model), but also in this case the estimated reaction rate and the estimated concentration of B suffer from noise. Also we are not able to explicitly tune the behavior of the estimated concentration of B, since it is assumed that the noise enters via v_r. In the previous chapter we were able to relax the noise behavior in the estimated concentration of B by increasing the noise $\bar{v}_{x_3}$. The chosen noise model suffers in this case from the lack of this possibility.

4.3.4. Method 1, advanced model

Utilizing increased 'process knowledge' as given by (90), estimating k_0 instead of Q_r, the model is as given by (104)–(108) with the state vector as by (91). Further we erroneously assuming the standard approach that all the states experience additive noise, and no noise in the auxiliary parameters. The initial values are

$$\hat{x}_0 = [100 \quad 0.5 \quad 0 \quad 300 \quad 0.95k_0]^T \tag{118}$$

$$\text{Sigma points lower bounds}: \; [0, \quad 0, \quad 0, \quad 0, \quad 8.5 \cdot 10^{11}]^T$$
$$\text{Sigma points upper bounds}: \; [\infty, \quad \infty, \quad \infty, \quad \infty, \quad 9.5 \cdot 10^{11}]^T \tag{119}$$

$$P_0 = diag[\, (10^{-5})^2 \quad (10^{-5})^2 \quad (10^{-5})^2 \quad (10^{-5})^2 \quad (9.45 \cdot 10^{10})^2 \,] \tag{120}$$

$$Qk = diag[\, \bar{v}_{k,x_1} \quad \bar{v}_{k,x_2} \quad \bar{v}_{k,x_3} \quad \bar{v}_{k,x_4} \quad \bar{v}_{k,x_5} \,]$$
$$= diag[\, (10^{-5})^2 \quad (10^{-5})^2 \quad (10^{-5})^2 \quad (10^{-5})^2 \quad (10^{11})^2 \,] \tag{121}$$

$$R_k = \bar{w}_k^2 = 0.075^2 \tag{122}$$

The results are as given by Fig. 14.

As is seen by Fig. 14, the estimation of all the states are close to the true ones, and the estimator also converges to the true state with respect to the estimated parameter k_0. However, we experienced that in the attempt to increase the convergence speed with respect to the estimated parameter k_0 (by increasing $\bar{v}_{x_5}$), the estimate of k_0 became more and more oscillatory.

4.3.5. Methods 1 and 3 combined, advanced model

Utilizing extended 'process knowledge' as given by (90), estimating k_0 instead of Q_r, the model is as given by (104)–(108) with the state vector as by (91). G is as by (110) and GQG^T as by (111). It is assumed that the reaction rate (r) experience Gaussian noise and is given by

$$\hat{r} = \hat{k}_0 e^{\left(\frac{-ER}{x_4}\right)} \hat{x}_2 \hat{x}_3^2 + v_r \tag{123}$$

The state vector is as given by (88) and the other initial values are

$$\hat{x}_0 = [100 \quad 0.5 \quad 0 \quad 300 \quad 0.95k_0]^T \tag{124}$$

$$\text{Sigma points lower bounds}: \; [0, \quad 0, \quad 0, \quad 0, \quad 8.5 \cdot 10^{11}]^T$$
$$\text{Sigma points upper bounds}: \; [\infty, \quad \infty, \quad \infty, \quad \infty, \quad 9.5 \cdot 10^{11}]^T \tag{125}$$

$$P_0 = diag[\, (10^{-5})^2 \quad (10^{-5})^2 \quad (10^{-5})^2 \quad (10^{-5})^2 \quad (9.45 \cdot 10^{10})^2 \,] \tag{126}$$

$$Q_k = diag[\, \bar{v}_{k,x_1} \quad \bar{v}_{k,r} \quad \bar{v}_{k,x_5} \,] = diag[\, (10^{-5})^2 \quad (10^{-3})^2 \quad (1.1 \cdot 10^{11})^2 \,] \tag{127}$$

$$R_k = \bar{w}_k^2 = 0.1^2 \tag{128}$$

The results are as given by Fig. 15.

As is seen by Fig. 15, the estimation of all the states are close to the true ones, and the estimator also converges to the true state with respect to the estimated parameter k_0. However, we experi-
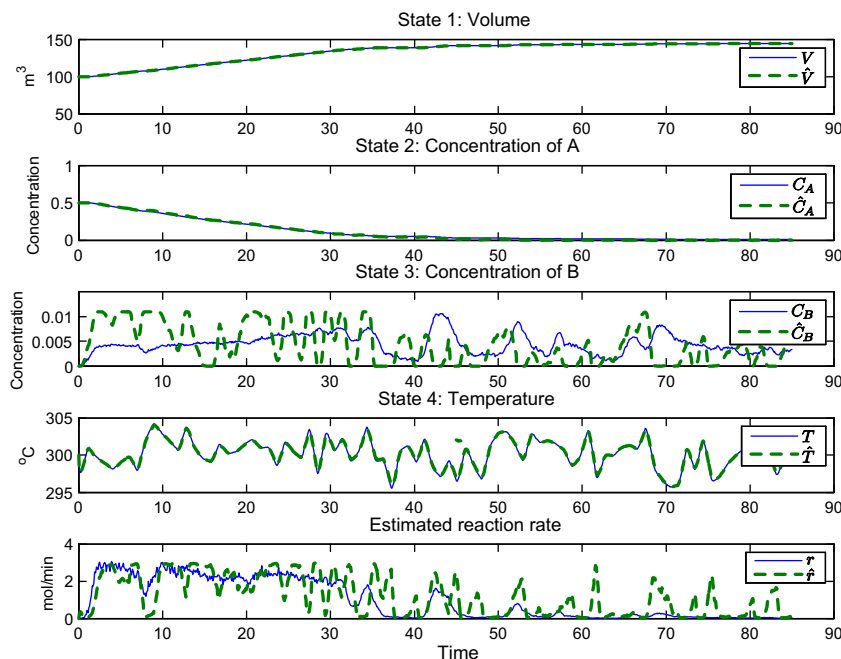

Fig. 13. The results of the simulation assumed simplified model with noise entering the reaction rate $\hat{r}$.
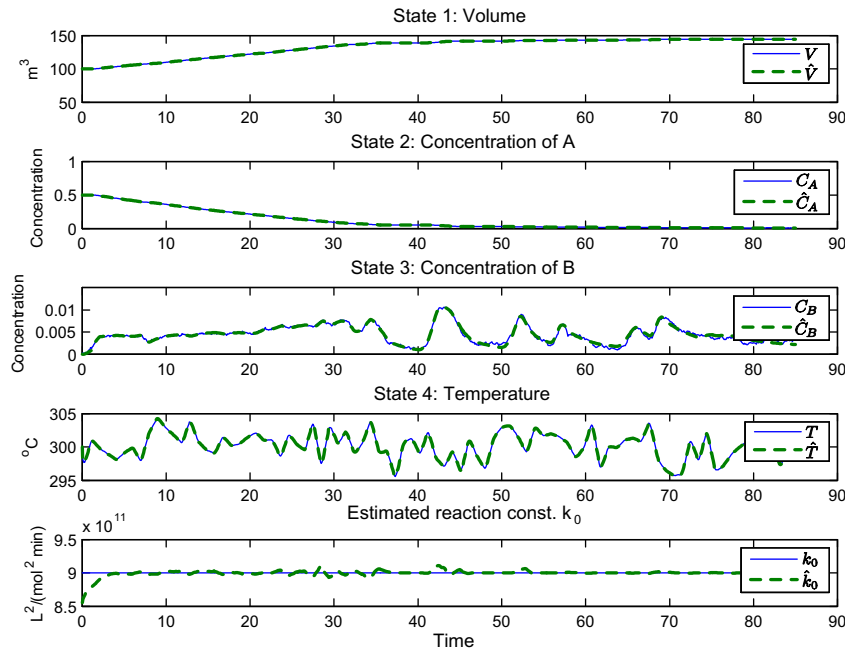
Fig. 14. The results of the simulation assumed increased process knowledge and assuming additive noise.
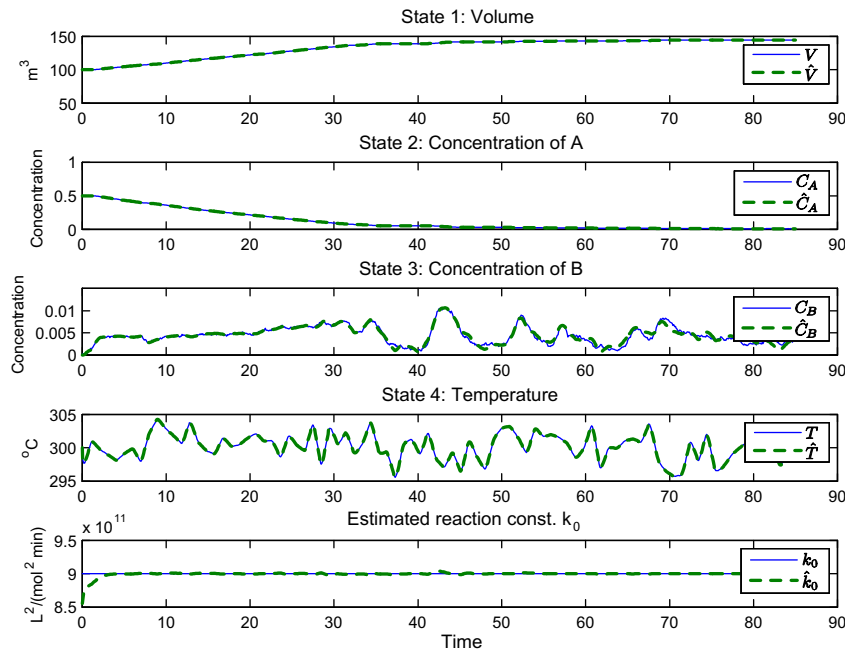

Fig. 15. The results of the simulation assumed increased process knowledge and assuming noise entering the reaction rate $\hat{r}$.

enced that in the attempt to further increase the convergence speed with respect to the estimated parameter k_0 (by increasing v_{x_5}), the estimate of k_0 become more and more oscillatory, but not as oscillatory as in the case of additive noise.

4.3.6. Some notes regarding Case '4 State CSTR'

By comparing our results with the results from Rao [20] we see that our results differs. This is mainly because we have decreased the noise variance of the temperature measurement from 1 to 0.1. This can be justified by that measuring the temperature with 0.1 °C accuracy is not unusual. This is especially of great impor-

tance using the simplified model, but of minor importance using the advanced model.

5. Discussion

The results indicate that the noise modeling concepts based on input uncertainty has an effect. In both the '2 state CSTR' case and the industrial case the same modeling concept (combination of Methods 1 and 2) is applied. However, the two simulated cases may be interpreted differently.

In the case of the '2 state CSTR' the purpose was to utilize 'process knowledge' that the states were uncertain when the control inputs were applied to the process. This is reflected by injecting the uncertainty into the covariance. The common way to handle such a situation, if accounted for at all, is believed to be to reset the covariance $P(t)$ to its initial value $P(0)$.

In the industrial case there was actual uncertainty in the control input, and the purpose was indeed to let this knowledge be incorporated into the covariance. The traditional way to handle such a situation is to apply a parameter to the control input, say $k_1(t)u_1(t)$ and try to estimate k_1 (see e.g. [19]). The presented method may be seen as an alternative to the traditional way, but without the need to estimate the parameter k_1, and hence avoid the need for state vector augmentation.

The investigation of the noise entering auxiliary variables in the model, as covered by Method 3, shows that it might serve as an alternative method to explicit off-diagonal tuning of the 'system noise matrix' Q/Q_k in cases where off-diagonal tuning is considered an option. Further, this method could be seen as a way not to jeopardize the mass/energy balance, in that the correlation between the states when it comes to how the uncertainty enters the states is fulfilled. This is seen by e.g. (46) and (47).

The investigation of the '4 state CSTR'-case showed that all the investigated noise models fulfilled the main challenge, namely to estimate the concentration of A. In the extended challenge, to accurately also estimate the reaction rate and the concentration of B, the choice of noise models made a significant difference. In the case of an simplified model, Method 3 did not improve the remaining estimates. In the case of applying a more accurate model, the use of Method 3 (combined with Method 1) showed marginally better estimation properties with respect to the estimated parameter k_0. Also, reducing the noise vector, implies fewer tuning parameters and noise parameters to be determined, which made the approach in the said case easier to tune than the additive noise approach.

6. Conclusion

We have presented some alternative noise modeling concepts which may serve as an effective alternative to the 'additive noise' paradigm.

Further, we have shown that letting noise entering auxiliary variables, it might serve as an alternative method to explicit off-diagonal tuning of the 'system noise matrix' Q/Q_k in cases where off-diagonal tuning is considered an option.

Acknowledgements

References

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

[2] D.F. Bizup, D.E. Brown, The over-extended Kalman filter – don't use it!, in: Proceedings of the Sixth International Conference of, vol. 1, 2003, pp. 40–46.

[3] W. Chen, L. Lang, B.R. Bakshi, P.K. Goel, S. Ungarla, Bayesian Estimation via Sequential Monte Carlo Sampling – Constrained Dynamic Systems, Automatica 9 (43) (2007) 1615–1622.

[4] T. Drengstig, D. Ljungquist, B. Foss, On the AlF$_3$ and temperature control of an aluminum electrolysis cell, IEEE Transactions on Control Systems Technology 6 (1998) 157–171.

[5] B. Foss, T.S. Schei, Putting Nonlinear Model Predictive Control into Use, Assessment and Future Directions of Nonlinear Model Predictive Control, 358, Springer-Verlag, 2007. pp. 407–417.

[6] A. Gelb, Applied Optimal Estimation, The MIT Press, ISBN 0-262- 20027-9, 1974.

[7] E. Gran, A multi-variable control in aluminum reduction cells, Modeling Identification and Control 1 (4) (1980) 247–258.

[8] K. Grjotheim, H. Kvande, Introduction to aluminium electrolysis, Aluminum-Verlag, 1993.

[9] E.L. Hasseltine, J.B. Rawlings, A Critical evaluation of Extended Kalman Filtering and Moving Horizon Estimation. TWMCC, Technical report number 2002-03, 2003.

[10] S.J. Julier, The scaled unscented transform, Proceedings of the American Control Conference 6 (2002) 4555–4559.

[11] S.J. Julier, J.K. Uhlmann, A general method for approximating nonlinear transformation of probability distributions, Technical report, RRG, Department of Engineering Science, University of Oxford, Can be downloaded from <http://www.robots.ox.ac.uk/~siju/work/publications/lettersize/Unscented.zip>, August 1994.

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

[13] S.J. Julier, J.K. Uhlmann, H. Durrant-Whyte, A new method for the nonlinear transformation of means and covariances in filters and estimators, IEEE Transactions on Automatic Control 45 (3) (2000) 477–482.

[14] R. Kalman, A new approach to linear filtering and prediction problems, Transaction of the ASME – Journal of Basic Engineering 82 (Series D) (1960) 35–45.

[15] R. Kandepu, B. Foss, L. Imsland, Applying the unscented Kalman filter for nonlinear state estimation, Journal of Process Control 18 (2008) 753–768.

[16] S. Kolås, B. Foss, T.S. Schei, Constrained nonlinear state estimation based on the UKF approach, Computers and Chemical Engineering, accepted for publication.

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

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

[19] G. Pannoncchia, Robust disturbance modeling for model predictive control with application to multivariable ill-conditioned processes, Journal of Process Control 13 (8) (2003) 693–701.

[20] C.V. Rao, Moving Horizon Strategies for the Constrained Monitoring and Control of Nonlinear Discrete-Time Systems, Ph.D. Thesis (Chemical Engineering), University of Wisconsin-Madison, 2000.

[21] J.B. Rawlings, N.F. Jerome, J.W. Hamer, T.M. Bruemmer, End-point control in semi-batch chemical reactors, in: Proceedings of the IFAC Symposium on Dynamics and Control of Chemical Reactors, Distillation Columns and Batch Processes, 1989, pp. 323–328.

[22] T. Saksvikrønning, E. Gran, K. Vee, Estimation of States in Aluminum Reduction Cells Applying Extended Kalman Filtering Algorithm Together with a Nonlinear Dynamic Model and Direct Measurements, TMS, AIME-meeting, Las Vegas, USA, 1976.

[23] T.S. Schei, A finite-difference method for linearization in nonlinear algorithms, Automatica 33 (11) (1997) 2053–2058.

[24] H. Schuler, C.U. Schmidt, Calorimetric state estimators for chemical reactor diagnosis and control: review of methods and applications, Chemical Engineering Science 47 (4) (1992) 899–915.

[25] D. Simon, Optimal State Estimation, Wiley, ISBN-10 0-471-70858-5, 2006.

[26] T. Söderstrøm, P. Stoica, System Identification, Prentice-Hall, 1989.

[27] P. Vachhani, S. Narasimhan, R. Rengaswamy, Robust and reliable estimation via unscented recursive nonlinear dynamic data reconciliation, Journal of Process Control 16 (2006) 1075–1086.

[28] P. Vachhani, R. Rengaswamy, V. Gangwal, S. Narasimhan, Recursive estimation in constrained nonlinear dynamical systems, AIChE Journal 51 (3) (2005) 946–959.

[29] R. van der Merwe, Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models, Ph.D. Dissertation, OGI School of Science and Engineering at Oregon Health and Science University, 2004.