Nonlinear Stochastic Modeling to Improve State Estimation in Process Monitoring and Control

Fernando V. Lima and James B. Rawlings
Dept. of Chemical and Biological Engineering, University of Wisconsin–Madison, Madison, WI 53706

DOI 10.1002/aic.12308
Published online June 25, 2010 in Wiley Online Library (wileyonlinelibrary.com).

State estimation from plant measurements plays an important role in advanced monitoring and control technologies, especially for chemical processes with nonlinear dynamics and significant levels of process and sensor noise. Several types of state estimators have been shown to provide high-quality estimates that are robust to significant process disturbances and model errors. These estimators require a dynamic model of the process, including the statistics of the stochastic disturbances affecting the states and measurements. The goal of this article is to introduce a design method for nonlinear state estimation including the following steps: (i) nonlinear process model selection, (ii) stochastic disturbance model selection, (iii) covariance identification from operating data, and (iv) estimator selection and implementation. Results on the implementation of this design method in nonlinear examples (CSTR and large dimensional polymerization process) show that the linear time-varying autocovariance least-squares technique accurately estimates the noise covariances for the examples analyzed, providing a good set of such covariances for the state estimators implemented. On the estimation implementation, a case study of a chemical reactor demonstrates the better capabilities of MHE when compared with the extended Kalman filter. © 2010 American Institute of Chemical Engineers *AIChE J,* 57: 996–1007, 2011
Keywords: state estimation, process control, nonlinear modeling, stochastic modeling, covariance estimation

Introduction and Prior Work

State estimation from plant measurements plays a critical role in all advanced monitoring and control systems. The state estimator is the component of the system that assesses in realtime the feedback from the measurements, separates the noise from the signal, and infers based on a dynamic model what information the sensors are providing about the current state of the system. This task is challenging because many of the best available models for chemical processes are nonlinear and large dimensional, the product properties of most interest are complex nonlinear functions of the process state, the structure of the stochastic disturbances affecting the process is unknown a priori, and the quality of the measurements is not uniform. This article introduces a state estimation design procedure to address this challenging and industrially important class of problems.

This procedure has several steps from nonlinear stochastic modeling to estimator implementation. Regarding the modeling task, the use of linearized models to represent a nonlinear chemical process is a common practice in process identification, estimation, and control. However, these models do not accurately describe the dynamics of the nonlinear process. The modeling approach proposed here relies on

Correspondence concerning this article should be addressed to F. V. Lima at flima@bevo.che.wisc.edu.

combining information that is often available such as a deterministic set of nonlinear differential equations describing the physical principles of the process, which arise from conservation laws appropriate for chemical processes, and a routine set of operating data that provide a typical sample of the measurement and process disturbances affecting the system. Integrating disturbance models are proposed to provide offset free control of the properties of interest, while maintaining the complexity low enough so that the disturbance statistics can be determined from the available measurements. The developed models will be used to improve nonlinear state estimation (e.g., MHE) in process monitoring and advanced process control.

Concerning the stochastic component of the model, the identification of the noise statistics affecting the states and the measurements will be performed with the assumption that these noises can be modeled as zero mean Gaussian random variables, and thus only their covariances are required to specify their statistics. These statistics are usually unknown, but can be estimated from process operating data. The autocovariance least-squares (ALS) technique has proven to accurately estimate the covariances of system disturbances from data for linear[1–4] and nonlinear systems,[5] using both simulated data and process operating data. This technique uses routine process operating data, and thus does not require input-output testing to be applied to the system. Simply stated, autocorrelations of the data at different time lags are taken to separately estimate noise covariances affecting the states and the measurements. Specifically, the linear time-varying ALS (LTV-ALS) technique was developed to estimate these covariances for nonlinear systems. Despite of the use of linearizations, this approach has accurately estimated the covariances of a nonlinear blending drum industrial example[5] (see Section Time-varying Autocovariance Least-Squares Technique for a mathematical summary of the LTV-ALS technique formulated for nonlinear models). Other correlation-based approaches to estimate noise covariances for linear systems can be found in Refs. 6–12. Also, for nonlinear systems, approaches based on maximum likelihood estimation (MLE), that still use knowledge from first-principles models, were developed.[13–17] The main drawbacks of these nonlinear covariance estimation approaches are that they are unconstrained, the optimization algorithms they use are iterative, and convergence to the optimal solution is not guaranteed. As a consequence, the estimation of covariances are highly dependent on the initial guesses and positive definite results are not guaranteed. Therefore, as a condition to obtain accurate estimates, these approaches may require the use of large data sets. However, MLE techniques are also highly dependent on the size of the data set used and may become computationally intractable for relevant industrial data sets. Rajamani[18] (Chapter 6) provides a comparison between ALS and MLE. In this comparison, the author concluded that the MLE procedure could not be applied to industrial data sets due to the large memory requirements even for problems of small dimensions. Finally, Valappil and Georgakis[19] introduced two approaches for estimation of noise statistics for nonlinear systems, one based on linearization and the other based on Monte Carlo simulations. However, these approaches assume that the measurement covariance matrix is known and can be determined independently for the measurement device, and hence does not take into account the trade-off between process and measurement noise covariances during the estimation. Moreover, these methods require information about the process–model mismatch in the form of a parameter covariance matrix. The authors mention that in case the users do not have such information, different covariance matrices for the parameters should be attempted. This contrasts with the ALS approach that estimates the stochastic structure of the disturbance model from data, automatically rescaling its estimated covariances to take such mismatches into account.

Regarding the nonlinear state estimation task, several state estimation techniques are available in the literature, divided into three main categories. These categories, along with the most popular estimators in each category, are the following: (1) purely recursive: extended Kalman filter (EKF),[20] unscented Kalman filter (UKF),[21–25] particle filter (PF),[26–31] and cell filter (CF),[32,33] (2) optimization-based: Moving Horizon Estimation (MHE),[34–38] and (3) hybrid methods combining (1) and (2): PF-MHE.[39,40] Rawlings and Bakshi[39] present an overview of many of these methods. Purely recursive state estimators may not be robust to the presence of data outliers, process disturbances and model errors. The optimization-based estimators use a sliding window of measurements that provides some inherent robustness to these errors, but may become computationally prohibitive when applied to large-scale industrial applications because of the requirement of solving a nonlinear programming problem online. Moreover, the optimization-based methods handle nonlinear models without linearization and incorporate hard state constraints, such as nonnegative concentrations and pressures, in the problem formulation.[34–37] For some applications, the use of hybrid approaches that combine the speed of recursive methods and the robustness of optimization-based methods may be an attractive alternative.[39] Finally, with increasing computer power and the development of faster algorithms to solve the MHE optimization problem online, pure MHE may become suitable for even large-scale applications.[41–44] Specifically, Zavala et al. have proposed fast MHE strategies exploiting the recent advances in nonlinear programming algorithms and sensitivity calculations.[41,42]

Another objective of this article is to make a critical assessment of some of the existing covariance and state estimation tools through their application to nonlinear case studies. The outline of the rest of the article is as follows. First, the modeling issues are presented and we show that a continuous time (CT) nonlinear stochastic differential equation (SDE) system model can be well represented by a discrete time (DT) nonlinear deterministic model, obtained from the integration of a first-principles model, plus an added noise component that can be estimated from operating data containing the process and measurement noise. Also, it is demonstrated that the process noises used in such DT models are well approximated by normal distributions with time invariant statistics. Then, the covariances of the process (Q) and measurement (R) noises are estimated using the LTV-ALS technique formulated for nonlinear models[5]. To perform this covariance estimation step, two nonlinear process examples are considered: (1) a continuous stirred-tank reactor (CSTR) example from[45] and (2) a published model of a gas-phase

ethylene copolymerization process.[46–48] These estimated noise covariances are used to specify the noise statistics of MHE and EKF state estimators. Using the illustrative CSTR example, these estimators are implemented and compared to show that high-quality state estimates can be obtained after the determination of their noise statistics by ALS. Finally, a summary of the proposed design method is presented followed by the conclusions and future research directions.

Nonlinear Stochastic Model Structure

We show here that a nonlinear stochastic model for the states in CT can be well represented by a nonlinear deterministic model in DT, obtained from the integration of a first-principles model, with an added noise component that can be estimated from process operating data, which is assumed available at every process sampling time $\Delta_k = t_{k+1} - t_k$. Assume the plant model is given by the following set of CT nonlinear stochastic differential equations (SDEs) for the states and nonlinear equations for the measurements [13]

$$dx = f(x,u)dt + \sigma dw \tag{1}$$

$$y_k = h(x_k) + v_k \tag{2}$$

in which $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, and $y \in \mathbb{R}^p$ are the states, manipulated inputs, and measured outputs, respectively. Also, w are the process noises, represented by standard Wiener processes, and $v_k \in \mathbb{R}^p$ are the measurement noises, which are assumed to be Gaussian with mean zero and covariance matrix R ($v_k \sim N(0,R)$).

This SDE can be solved at time t_T using the Euler scheme[49] that consists of the following discrete approximation of the SDE at time t_T

$$x_{T+1} = x_T + f(x_T,u_T)\Delta_T + \sigma\Delta w_T \tag{3}$$

in which Δ_T is the Euler integration stepsize that is chosen small enough to provide an accurate SDE solution and to guarantee convergence of the Euler method ($\Delta_T \ll \Delta_k$). Also, $\Delta w_T = w_{T+1} - w_T$ are normally distributed increments of the standard Wiener process.

For one step ahead predictions, we show next that model (1,2) can be well represented by the following DT model

$$x_{k+1} = F(x_k,u_k) + G(x_k)w_k \tag{4}$$

$$y_k = h(x_k) + v_k \tag{5}$$

in which $F(x_k,u_k)$ is obtained by integrating the deterministic nonlinear model $f(x,u)$ from t_k to t_{k+1} using an ordinary differential equation (ODE) solver with a zero-order hold on the input u_k, and $w_k \sim N(0,Q)$. For purposes of illustration, consider the simplified example

$$\text{SDE}: x_{T+1} = x_T + f(x_T)\Delta_T + \Delta w_T \tag{6}$$

$$\text{DT model}: x_{k+1} = F(x_k) + w_k \tag{7}$$

in which the output equations, that are discrete for both models (2) and (5), are omitted and the input variables dropped for simplicity. These variables can be easily incorporated in the analysis that follows.

To determine the noise statistics of w_k, assume the simulated data generated by the SDE solution (6) represent the plant data.[13] These plant data are discretized using the process sampling time Δ_k. Thus, using (7), an estimate of w_k, denoted by $\hat{w}_k$, can be calculated by

$$\hat{w}_k = \underbrace{x_{k+1}}_{\text{from SDE solution}} - \underbrace{F(x_k)}_{\text{from ODE solver}} \tag{8}$$

in which x_{k+1} and x_k are obtained from the discretized data points generated by the SDE solution, or plant data. Note that $\hat{w}_k$ corresponds to the residuals between the SDE and the deterministic DT solutions for the same x_k. Thus, if several simulations with different realizations of noise sequences are performed one step ahead at time t_k, then the statistics of w_k (sample mean and variance) can be estimated using all the values of $\hat{w}_k$ obtained for all these simulations. Once these statistics are determined, (7) can be used as the model in state estimation and model-based control. In the next subsection, a case study is performed to show that these statistics are well approximated as normal and time invariant.

Illustrative batch reactor example

Consider a nonlinear, chemical batch reactor[45] consisting of the following set of reversible reactions in the gas-phase

$$\text{A} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \text{B} + \text{C} \quad 2\text{B} \underset{k_4}{\overset{k_3}{\rightleftharpoons}} \text{C}$$

and the nonlinear model from first principles

$$\frac{dx}{dt} = f(x) = \nu' r \tag{9}$$

in which the state is $x = [c_A \; c_B \; c_C]'$ and the stoichiometric matrix, reaction rates, and parameters are given by

$$\nu = \begin{bmatrix} -1 & 1 & 1 \\ 0 & -2 & 1 \end{bmatrix}, \qquad r = \begin{bmatrix} k_1 c_A - k_2 c_B c_C \\ k_3 c_B^2 - k_4 c_C \end{bmatrix},$$

$$k = [\,0.5 \quad 0.05 \quad 0.2 \quad 0.01\,]', \qquad x_0 = [\,0.5 \quad 0.05 \quad 0\,]'$$

Let Δw denote the vector of random variables added to the species concentrations. We assume Δw is normally distributed with zero mean and the covariance Q_{sim} below ($\Delta w \sim N(0,Q_{\text{sim}})$). The solution of the SDE, (6), for this example using the Euler scheme is given by

$$\begin{bmatrix} c_A \\ c_B \\ c_C \end{bmatrix}_{T+1} = \begin{bmatrix} c_A \\ c_B \\ c_C \end{bmatrix}_T + \nu' r(c)\Delta_T + \begin{bmatrix} \Delta w_1 \\ \Delta w_2 \\ \Delta w_3 \end{bmatrix}_T \tag{10}$$

$$Q_{\text{sim}} = 10^{-7} \times \begin{bmatrix} 1.25 & 0.00 & 0.00 \\ 0.00 & 1.25 & 0.00 \\ 0.00 & 0.00 & 1.25 \end{bmatrix}$$

Thus, for this case study, Eq. 10 is solved recursively to generate the simulated data, (9) is used for the deterministic

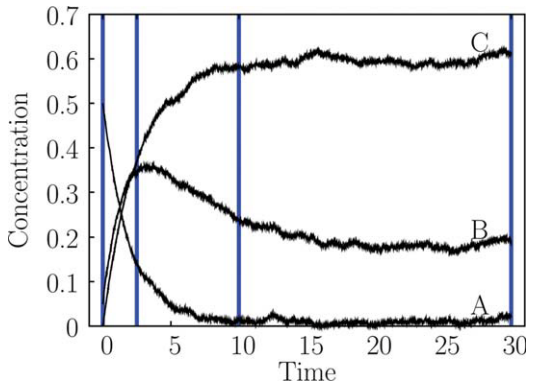

Figure 1. Evolution of the concentrations for a typical SDE simulation (dark line).

The time instants from where one step ahead simulations are performed are represented by the thick vertical lines. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

first principles model, and the following step sizes for SDE simulation and measurement sample time are chosen

$$\Delta_T = 0.25 \times 10^{-2}, \qquad \Delta_k = 0.25$$

Figure 1 shows the concentrations of A, B, and C for a typical SDE simulation from an initial condition of almost pure A until the system reaches steady state. To calculate the statistics of $\hat{w}_k$, four cases with different starting states are considered: initial condition: $x(t_k = 0) = x_0$, two transient cases: $x(t_k = 2.5) = [0.149\ 0.345\ 0.379]'$ and $x(t_k = 10) = [0.02\ 0.257\ 0.617]'$, and a steady-state case $x(t_k = 30) = [0.012\ 0.186\ 0.663]'$. These four cases represent different operating conditions of the reactor process and are displayed in Figure 1 by the thick vertical lines.

Figure 2 shows the histograms of $\hat{w}_1$, $\hat{w}_2$, and $\hat{w}_3$ when one step ahead simulations are performed from $t_k = 0$. These histograms are obtained using all the values of $\hat{w}_k$ calculated for 1000 simulations with different realizations of noise sequences. Note that the components of $\hat{w}_k$ are normally distributed with zero mean and the following multivariable sample variance

$$\hat{Q}_0 = 10^{-5} \times \begin{bmatrix} 1.09 & 0.05 & 0.09 \\ 0.05 & 1.22 & -0.03 \\ 0.09 & -0.03 & 1.17 \end{bmatrix}$$

in which the subscript "0" denotes that this covariance was estimated for the $x(t_k = 0)$ case. The same analysis is repeated for the other cases when one step ahead predictions are made from $x(t_k = 2.5)$, $x(t_k = 10)$, $x(t_k = 30)$ and the following covariances are obtained

$$\hat{Q}_{2.5} = 10^{-5} \times \begin{bmatrix} 1.09 & 0.05 & 0.09 \\ 0.05 & 1.16 & -0.01 \\ 0.09 & -0.01 & 1.16 \end{bmatrix}$$

$$\hat{Q}_{10} = 10^{-5} \times \begin{bmatrix} 1.09 & 0.06 & 0.09 \\ 0.06 & 1.18 & -0.02 \\ 0.09 & -0.02 & 1.16 \end{bmatrix}$$

$$\hat{Q}_{30} = 10^{-5} \times \begin{bmatrix} 1.09 & 0.06 & 0.09 \\ 0.06 & 1.19 & -0.03 \\ 0.09 & -0.03 & 1.17 \end{bmatrix}$$

with approximately zero mean for all cases. Note that the calculated variances of $\hat{w}_k$ are close to each other regardless of the initial state considered. Thus, this example demonstrates that not only are the $\hat{w}_k$ normally distributed, but their statistics are time invariant even while the process state is changing significantly. The normality of the DT model process noise is a consequence of the central limit theorem that states that the

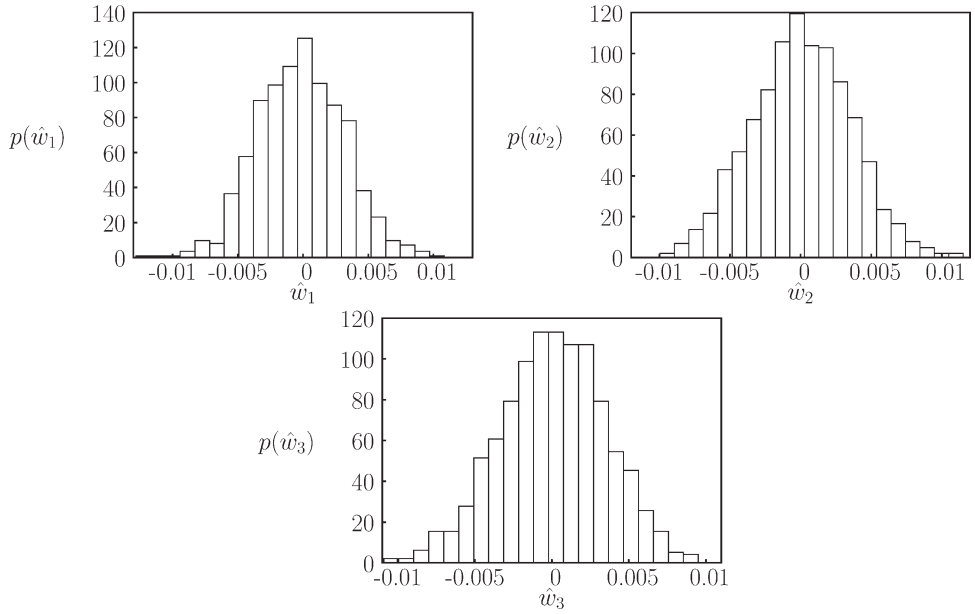

Figure 2. Results for $x(t_k = 0)$ case: $\hat{w}$ histograms for 1000 simulations.

addition of i independent and identically distributed random variables with density p_i tends to a normal curve as i increases, regardless of the shape of the original densities p_i. The fact that the SDE is sampled 100 times faster than the DT model (Δ_k/Δ_T) helps to build up the number of added noises for the DT model, resulting in normal distributions. Therefore, adding a process noise with these characteristics to the integrated deterministic ODE as in (4) results in an accurate nonlinear model for one step ahead predictions. This model could also be used, but would not be as accurate, for multistep ahead predictions.

In summary, even with the assumption that SDE (1) is the plant generating the data, it is possible to obtain a highly accurate model of the data for one step ahead predictions using (4).

Noise Covariance Estimation for Nonlinear Systems

With the chosen nonlinear model and noise structure, we next turn to the task of determining the process and output noise statistics given only routine process output measurements. In this article, we consider a method based on linearizations of the nonlinear model to estimate the disturbance variances.

Time-varying autocovariance least-squares technique

Here, we present a summary of the linear time-varying autocovariance least-squares (LTV-ALS) technique for nonlinear systems originally proposed by Rajamani et al.[5] We use the following discrete time nonlinear model

$$\begin{aligned} x_{k+1} &= F(x_k, u_k) + G(x_k)w_k \\ y_k &= h(x_k) + v_k \end{aligned} \tag{11}$$

in which $w_k \sim N(0,Q)$ and $v_k \sim N(0,R)$ are uncorrelated with each other. Assume that a time-varying stable state estimator, such as the extended Kalman filter (EKF), with a stable filter gain sequence, L_k, is implemented to estimate the states of this model. In this case, the state estimation equations are given by

$$\begin{aligned} \hat{x}_{k+1|k} &= F(\hat{x}_{k|k}, u_k) \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + L_k(y_k - \hat{y}_{k|k-1}) \\ \hat{y}_{k|k-1} &= h(\hat{x}_{k|k-1}) \end{aligned} \tag{12}$$

in which $\hat{x}_{k|k-1}$ denotes the predicted estimate of the state x_k using the available information up to time t_{k-1}. Also, $\hat{x}_{k|k}$ represents the filtered estimate at time t_k.

If a linearization of nonlinear model (11) around $(\hat{x}_{k|k}, u_k)$ is performed, then this model can be represented by the following set of equations

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k + G_k w_k \\ y_k &= C_k x_k + v_k \end{aligned} \tag{13}$$

in which

$$A_k = \frac{\partial F(x_k,u_k)}{\partial x_k'}\Big|_{(\hat{x}_{k|k},u_k)}, \quad C_k = \frac{\partial h(x_k)}{\partial x_k'}\Big|_{\hat{x}_{k|k}}, \quad G_k = G(\hat{x}_{k|k}, u_k)$$

Also, the linearized estimation equations are given by

$$\begin{aligned} \hat{x}_{k+1|k} &= A_k\hat{x}_{k|k} + B_k u_k \\ \hat{y}_{k|k-1} &= C_k\hat{x}_{k|k-1} \end{aligned} \tag{14}$$

The innovations sequence at time t_k can be obtained by subtracting the output equation in (14) from that in (13)

$$\mathscr{y}_k = y_k - \hat{y}_{k|k-1} \tag{15}$$

Using this sequence at different time instants, the autocovariance matrix is defined as the expectation of the innovations data at different time lags over a user defined window N[50]

$$\mathscr{R}_k(N) = E\begin{bmatrix} \mathscr{y}_k\mathscr{y}_k' \\ \vdots \\ \mathscr{y}_{k+N-1}\mathscr{y}_k' \end{bmatrix} \tag{16}$$

An optimization problem of the following form can then be solved to calculate estimates of Q and R [see Ref. 5 for mathematical details on the construction of $\mathscr{A}_k$ and $\hat{b}_k$]

$$\begin{aligned} &\min_{Q,R} \left\| \mathscr{A}_k \begin{bmatrix} (Q)_s \\ (R)_s \end{bmatrix} - \hat{b}_k \right\|^2 \\ &\text{s.t.} \quad Q,R \geq 0, \quad Q = Q', \quad R = R' \end{aligned} \tag{17}$$

in which the subscript "s" denotes the column-wise stacking of the elements of a matrix into a vector. Also, the matrices $\mathscr{A}_k$ and $\hat{b}_k$ satisfy the following relationships

$$\mathscr{A}_k \begin{bmatrix} (Q)_s \\ (R)_s \end{bmatrix} = [\mathscr{R}_k(N)]_s, \quad \hat{b}_k = [\widehat{\mathscr{R}}_k(N)]_s$$

in which the innovations sequences used to calculate an estimate of the autocovariance matrix, denoted by $\widehat{\mathscr{R}}_k(N)$, are obtained using the plant data.

Thus, if a set of data with length N_d is available, then the following least-squares problem can be solved to estimate the covariances

$$\begin{aligned} &\min_{Q,R} \left\| \begin{bmatrix} \mathscr{A}_k \\ \vdots \\ \mathscr{A}_{N_d-N+1} \end{bmatrix} \begin{bmatrix} (Q)_s \\ (R)_s \end{bmatrix} - \begin{bmatrix} \hat{b}_k \\ \vdots \\ \hat{b}_{N_d-N+1} \end{bmatrix} \right\|^2 \\ &\text{s.t.} \quad Q,R \geq 0, \quad Q = Q', \quad R = R' \end{aligned} \tag{18}$$

in which the time-varying matrices $\mathscr{A}_k$ and $\hat{b}_k$ are computed using a sliding window strategy.[5]

Remark 1. The only requirement for the use of the extended Kalman filter as the state estimator to obtain the ALS innovations sequences is that this estimator has to satisfy the following assumption: the time-varying filter gain sequence L_k used in (12) is such that, when used in the approximate linearization given by (13), it produces a sequence of $\overline{A}_k = (A_k - A_kL_kC_k)$ matrices such that the product $\left(\prod_{i=k-1}^{0}\overline{A}_i\right) \to 0$ as k increases. This assumption can be guaranteed under suitable detectability and

stabilizability conditions[51] and is necessary to fulfill the steady-state assumption of ALS.[1,4]

Remark 2. The probabilistic approach for state estimation is to maximize the conditional density of the state sequence given the measurement sequence

$$\max_{\{x_k\}} p(\{x_k\}|\{y_k\})$$

An equivalent least-squares problem when noises are Gaussian is given by

$$\min_{x_0,x_1\cdots x_k} \|x_0 - \bar{x}_0\|_{P^{-1}} + \sum_{i=0}^{k-1} w_i Q^{-1} w_i' + v_i R^{-1} v_i'$$

$$\text{st : nonlinear model } (11) \text{ and constraints}$$

in which (Q,R) are the covariances of the nonlinear model and the inverse of the weights of this least-squares problem. This least-squares problem is similar to the one solved by MHE, and thus estimated covariances can be used to specify the noise statistics of this state estimator.

Remark 3. As the values of the ALS design parameters, horizon length (N) and data set length (N_d), increase, the accuracy of the covariance estimation improves with the expense of a larger computational cost to perform such an estimation. These parameter values are selected according to the application. For example, for the polymerization process presented in the "Results" section, values of $N = 100$ and $N_d = 6000$ are large enough to provide good covariance estimates.

Note that this approach requires linearizations of the nonlinear model around the current state estimate and input values. This linearization may not be an issue for noise covariance estimation as shown in the Results Section and discussed in the Conclusions.

Diagonal ALS technique

For large dimensional applications, estimating only the diagonal elements of the covariance matrices Q and R may be an attractive alternative when the $\mathscr{A}_k$ matrices of the least-squares problem defined by (18) are ill-conditioned. This estimation is also useful to increase the speed of the ALS computations.[18]

Assume each of the $\mathscr{A}_k$ matrices is partitioned in two matrices

$$\mathscr{A}_k = [\,\mathscr{A}_{k1} \quad \mathscr{A}_{k2}\,]$$

in which $\mathscr{A}_{k1}$ and $\mathscr{A}_{k2}$ multiply $(Q)_s$ and $(R)_s$, respectively, as in (17). For this estimation, we assume Q and R diagonal and apply a linear transformation to each of the $\mathscr{A}_k$ partitions to remove the columns that are associated with the off-diagonal elements of the covariance matrices. Thus, in the diagonal ALS formulation, each of the $\mathscr{A}_k$ matrices is modified by right-multiplying $\mathscr{A}_{k1}$ and $\mathscr{A}_{k2}$ with $\mathscr{T}_n$ and $\mathscr{T}_p$, respectively. This formulation is as follows

$$\min_{Q,R} \left\| \left[\mathscr{A}_{k1}\mathscr{T}_n \quad \mathscr{A}_{k2}\mathscr{T}_p\right] \begin{bmatrix} (\mathrm{diag}(Q))_s \\ (\mathrm{diag}(R))_s \end{bmatrix} - \hat{b}_k \right\|^2$$

$$\text{s.t.} \quad Q, R \geq 0$$

in which

$$\mathscr{T}_n = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ & & 0_{n\times n} & & \\ 0 & 1 & 0 & \dots & 0 \\ & & 0_{n\times n} & & \\ 0 & 0 & 1 & \dots & 0 \\ & & 0_{n\times n} & & \\ & & \vdots & & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix},$$

$$\mathscr{T}_p = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ & & 0_{p\times p} & & \\ 0 & 1 & 0 & \dots & 0 \\ & & 0_{p\times p} & & \\ 0 & 0 & 1 & \dots & 0 \\ & & 0_{p\times p} & & \\ & & \vdots & & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

and the transformation matrices $\mathscr{T}_n$ and $\mathscr{T}_p$ have dimensionalities $\mathscr{T}_n \in \mathbb{R}^{n^2\times n}$ and $\mathscr{T}_p \in \mathbb{R}^{p^2\times p}$, respectively.

Results

In this section, we show the results of the application of the design method described above to two nonlinear examples: a CSTR and a polymerization process.

CSTR example: ALS case studies

As a starting point on the use of the LTV-ALS technique for covariance estimation, two case studies of a continuous stirred-tank reactor (CSTR) example, modified from,[45] with the same reaction stoichiometry and kinetics of the batch reactor from subsection Illustrative Batch Reactor Example are presented here. These cases are performed to verify the robustness of the ALS technique under the presence of process disturbances and plant-model mismatch.

Consider a well-mixed, isothermal CSTR model

$$\frac{dx}{dt} = f(x) = \frac{Q_f}{V_R} c_f - \frac{Q_0}{V_R} x + \nu' r \tag{19}$$

with selected parameters

$$k = [\,0.5 \quad 0.5 \quad 0.2 \quad 0.01\,]', \quad c_f = [\,0.5 \quad 0.05 \quad 0\,]',$$
$$x_0 = [\,0.5 \quad 0.05 \quad 0\,]', \qquad Q_f = Q_0 = 1, \qquad V_R = 100$$

To generate the simulated data for the first-case study, an oscillation of ±50% with a period of 100 time units was applied around the nominal value of the second reaction rate constant ($k_2 = 0.5 + 0.25\sin(2\pi t/100)$). Also, a zero mean white noise with variance $Q_{sim} = 1.25 \times 10^{-3}$ was added to this oscillating rate. In practice, these rate constant oscillations may occur due to temperature fluctuations in the reactor. Under such conditions, concentration data were generated by solving model (19) recursively at every Δ_T using the Euler method, as described in Section Nonlinear Stochastic

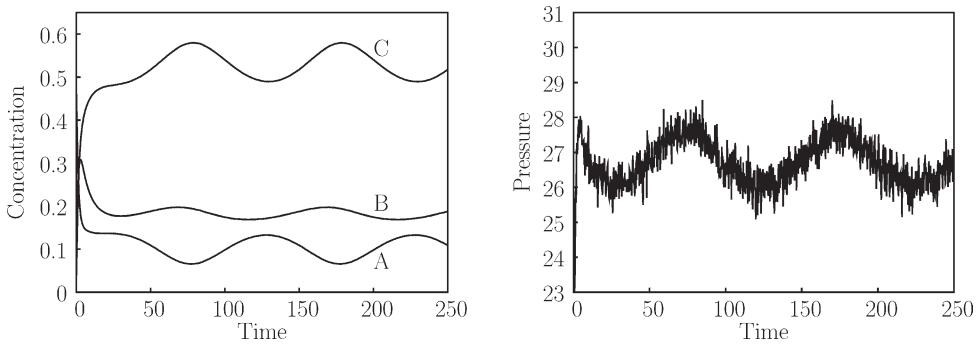

Figure 3. Typical simulated data set used for ALS calculations of Case 1.

Model Structure. Also, pressure data were generated at every Δ_T using the following equation

$$y_T = R_g T(c_A + c_B + c_C)_T + v_T$$

in which R_g is the ideal gas law constant and T is the reactor temperature. The measurement y_T denotes the pressure values at time t_T, $\Delta_T = 0.25 \times 10^{-2}$, and the noise sequence v_T is normally distributed with zero mean and variance $R_{sim} = 1.25 \times 10^{-1}$. To represent a plant data set, these simulated data for states and measurements were sampled with the process sampling time of $\Delta_k = 0.25$. Figure 3 shows typical sampled data sets for concentrations and pressure generated in this fashion. Using ALS, we estimate the process and measurement noise covariances (Q,R) of the DT model structure given by (11) from simulated data sets.

For ALS estimation, a general linear time-varying process model is proposed to estimate the noise covariances and cope with potential plant-model mismatches. For this purpose, we augment the state vector with one integrated white noise d. The first-principles version of this selected model is given by

$$\frac{d}{dt}\begin{bmatrix} x \\ d \end{bmatrix} = f(x,d) = \begin{bmatrix} \frac{Q_f}{V_R} c_f - \frac{Q_0}{V_R} x + \nu' \bar{r} \\ 0 \end{bmatrix}$$

in which

$$\bar{r} = \begin{bmatrix} k_1 c_A - (k_2 + d) c_B c_C \\ k_3 c_B^2 - k_4 c_C \end{bmatrix}$$

Thus, the discrete time version of this model is obtained by integrating $f(x,d)$ as described previously (see Sections Nonlinear Stochastic Model Structure and Time-varying Autocovariance Least-Squares Technique) and is the following

$$\begin{bmatrix} x \\ d \end{bmatrix}_{k+1} = \underbrace{\begin{bmatrix} A & G_d \\ 0_{1\times n} & 1 \end{bmatrix}_k}_{\overline{A}_k} \underbrace{\begin{bmatrix} x \\ d \end{bmatrix}_k}_{\overline{x}_k} + \underbrace{\begin{bmatrix} 0_{n\times 1} \\ 1 \end{bmatrix}}_{\overline{G}} w_k$$

$$y_k = \underbrace{[C \quad 0]}_{\overline{C}} \begin{bmatrix} x \\ d \end{bmatrix}_k + v_k$$

in which

$$\overline{A}_k = \left.\frac{\partial F(\overline{x}_k)}{\partial \overline{x}_k}\right|_{(\hat{\overline{x}}_{k|k})}, \quad C = [1 \quad 1 \quad 1] R_g T$$

and $F(\overline{x}_k)$ is obtained by integrating $f(x,d)$. Also, G_d represents the disturbance model or the relationship between the states and the disturbance. The only requirements on the selection of the augmented matrices $\overline{A}$, $\overline{C}$, and $\overline{G}$ for ALS estimation are that the pair $(\overline{A}, \overline{G})$ is stabilizable and $(\overline{A}, \overline{C})$ is detectable. Hence, the selection of the disturbance model must be checked to not violate these conditions.

Applying ALS to the generated data set, the following covariances were estimated

$$Q_{als} = 3.59 \times 10^{-4}, \quad R_{als} = 1.20 \times 10^{-1}$$

For this calculation, the first 160 data points (corresponding to the transient part of the data set from 0 to 40 time units) were removed to fulfill the steady-state assumption of ALS (see Ref. 1 for details). Also, this estimation was

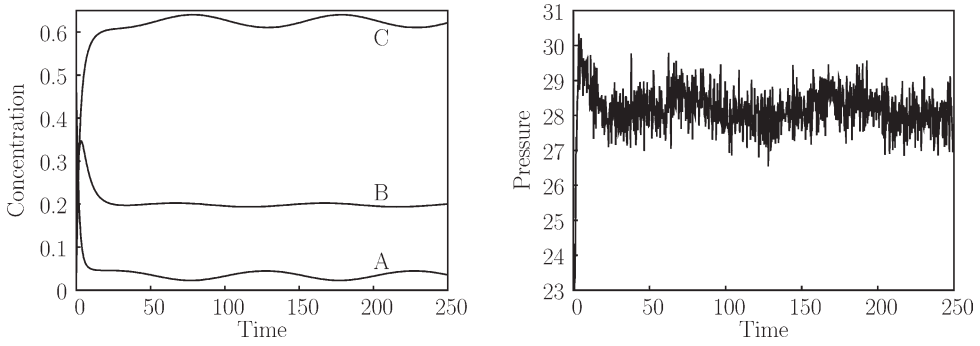

Figure 4. Typical simulated data set used for Case 2.

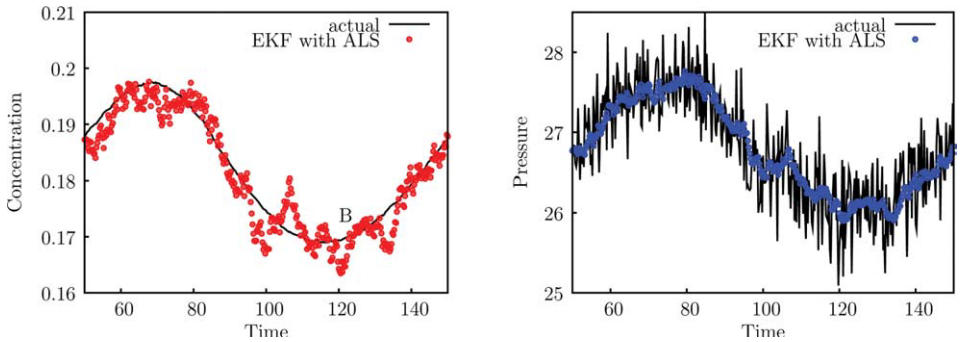

Figure 5. EKF implementation results for Case 1.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

repeated four times using different realizations of noise sequences, with parameters $N = 15$ and $N_d = 1000$, and the average of the estimated covariances for all these simulations was taken as the final result. Note that the estimated value of R_{als} is close to R_{sim}. The accuracy on the estimation of Q_{als} is verified in the next subsection where the estimated covariances are used to specify the state estimators.

As a second case study, assume in the true process, the measurement accuracy decreased over time and now $R_{sim} = 2.50 \times 10^{-1}$. Moreover, the rate constant k_2 was reduced due to the presence of a poison in the reactor and oscillates as $k_2 = 0.1 + 0.05 \sin(2\pi t/100)$. As in the previous case, a white noise was also added to this rate constant with a lower variance of $Q_{sim} = 1.25 \times 10^{-4}$. A new data set was generated for this case as shown in Figure 4, but ALS still uses the same process model as previously with the incorrect reaction rate constant. For this case, the ALS estimated covariances using the same procedure and parameters of Case 1 are

$$Q_{als} = 1.74 \times 10^{-5}, \quad R_{als} = 2.44 \times 10^{-1}$$

Note that, even though a wrong model was used for covariance estimation, ALS accurately calculates R and rescales Q as expected. In the next subsection, EKF and MHE, two widely used state estimators for nonlinear systems are implemented in these two case studies. Also, to confirm the accuracy of the ALS estimation, the covariances calculated here are used to specify the statistics of these state estimators.

CSTR: Estimator Selection and Implementation

Using the two case studies from the CSTR example of the previous subsection, here the implementation of EKF and MHE with statistics defined by the ALS covariances is performed.

For Case 1, Figure 5 shows a close look at the pressure and concentration of B for the EKF estimator using the ALS covariances and $x_0 = [0.51\ 0.05\ 0]'$. Note that EKF with ALS covariances follows both the true state and pressure measurement closely. Although only the plot for the concentration of B is shown here, the same holds for the concentration of the other components.

For Case 2 with $x_0 = [0.45\ 0.05\ 0]'$, however, the EKF with ALS covariances still does a good job estimating the pressure, but does not estimate the states accurately, with even some negative values for the concentration of A (see Figure 6). As shown in ref. 45, the EKF may not work well for nonlinear applications due to the linearization and the unconstrained nature of this estimator. The implementation of MHE, with statistics defined by the ALS covariances, is shown in Figure 7. Note that MHE overcomes the problem of negative A concentrations because of its capability to enforce hard state constraints. It is worth mentioning that priors slightly farther from the x_0 used to generated the data than the ones above would make the concentration of A take even more negative values and converge even slower to its steady-state. Some reasonable priors even lead to EKF divergence, as also reported previously in the literature.[45]

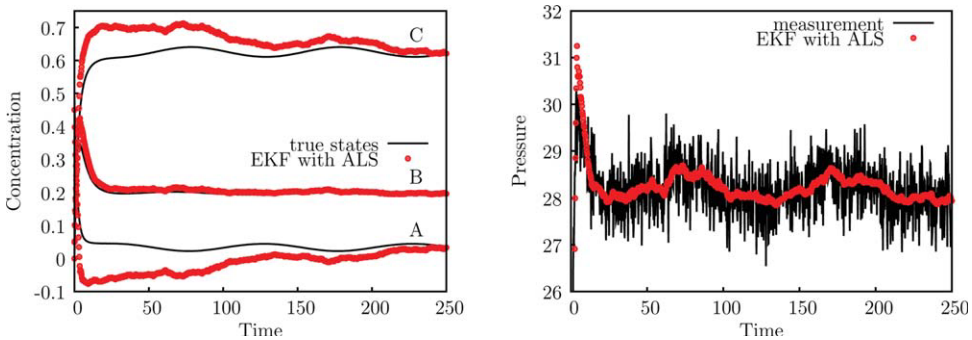

Figure 6. EKF implementation results for Case 2.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

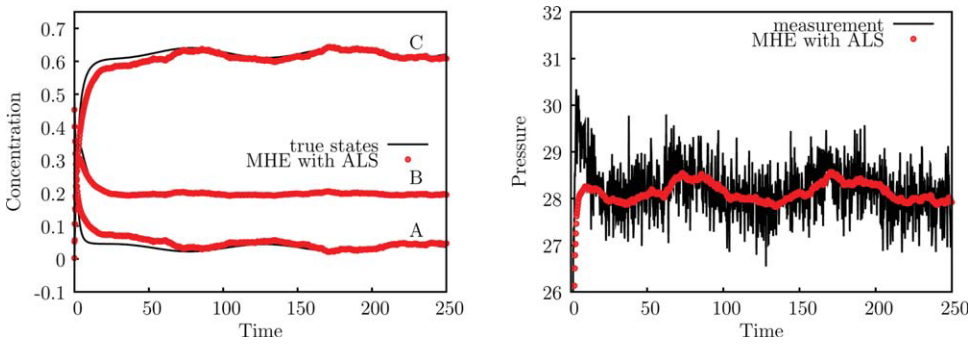

Figure 7. MHE implementation results for Case 2.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Polymerization Process

In this subsection, we present the results of the application of the described design method for nonlinear state estimation to a gas-phase ethylene copolymerization process model taken from the literature.[46–48] This polymerization process has the following variables: 41 states, 17 measurements, 5 inputs, process and measurement noises. Among the important variables of this process are the reactor temperature, pressure and compositions (ethylene, 1-butene (comonomer), inerts and hydrogen), production rate, and polymer properties (density and melting index). Figure 8 shows a schematic diagram of this process. See Refs. 46–48 for additional information about this process.

To model this polymerization system, a nonlinear first-principles model with an added noise component was developed, as explained in Section Nonlinear Stochastic Model Structure. This model was built based on the cited references above and has the structure of (4, 5). To generate the simulated data for this case study, Gaussian noises with assumed characteristics were added to the temperature of the recycle stream, representing a process noise, and to all the 17 measurements, representing sensor noises. The process sampling time is 60 s and the noise sequences $w_{\text{sim}} \sim N(0,Q_{\text{sim}})$ and $v_{\text{sim}} \sim N(0,R_{\text{sim}})$ have covariances

$$Q_{\text{sim}} = 2.80 \times 10^{-4}$$
$$R_{\text{sim}} = 10^{-6}\times$$
$$\text{diag}\,(5, 1, 10^4, 20, 300, 10, 200, 0.2,$$
$$0.5, 0.3, 0.04, 10^{-5}, 20, 1, 400, 300, 300)$$

Applying ALS, the estimated covariances of $w_k \sim N(0,Q)$ and $v_k \sim N(0,R)$ for the model structure (4, 5) are given by

$$Q_{\text{als}} = 2.84 \times 10^{-4}$$
$$R_{\text{als}} = 10^{-6}\times$$
$$\text{diag}(5.03, 1, 9.7 \times 10^3, 20.2, 308, 10.2, 198,$$
$$0.19, 0.5, 0.3, 0.04, 9.8 \times 10^{-6}, 27.2, 0.99, 419, 289, 296)$$

Note that once again the LTV-ALS technique estimates these covariances accurately. Figure 9 shows the temperature of the recycle streams (in K) plots for the EKF using the ALS covariances. However, it is worth noting that for the estimation of R, the diagonal ALS technique described in Section Diagonal ALS Technique was used to estimate only the diagonal components of this matrix. The information content in the data is too small to justify a more complex noise model structure. Thus, one relevant conclusion of this case study is that easily measured properties combined with a large dimensional state vector preclude the use of complex disturbance models for this system. We estimated the variance of a single process disturbance and only the diagonal components of the measurement covariance matrix. The user is provided with the information to detect this situation automatically by examining the conditioning of the $\mathscr{A}_k$ matrices of the least-squares problem defined by (18). An overly complex disturbance model for the available information in the measurements is detected by poor conditioning in the $\mathscr{A}_k$ matrices. In this situation, to reduce or eliminate the ill-conditioning problem that may also plague the state estimation step, approaches such as Schmidt-Kalman filtering may be attempted. This approach was originally developed for navigation systems to improve numerical stability and reduce computational complexity of Kalman filters by eliminating some of the state variables that are of no interest for the problem.[52,53] In this case, we intend to use this estimator to tackle weakly observable systems as described in Refs. 54 and 55. The general idea of this technique is to remove weakly observable states in the Kalman filter gain calculation, producing a suboptimal filter that does not estimate the removed state variables, but still keeps track of the

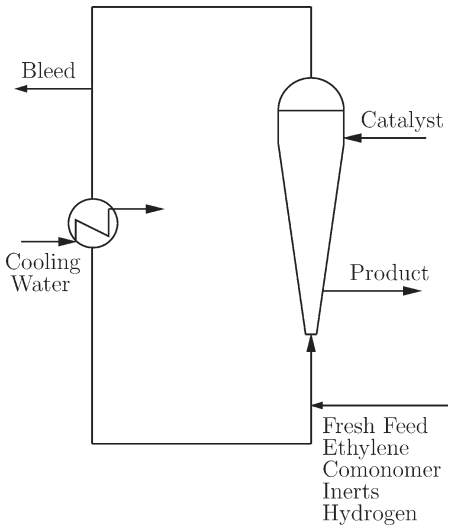

Figure 8. Gas-phase polyethylene reactor system.[48]

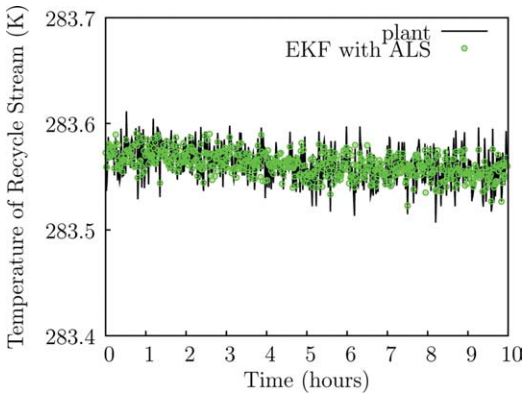

Figure 9. EKF implementation results for polymerization problem.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

influences these states have on the gain applied to the other states. The application of this technique to industrial processes is currently under investigation.

Summary of Design Method

This section summarizes the proposed design method for nonlinear state estimation. The steps of this method as well as the inputs, computations, and outputs associated with each step are the following:

(i) nonlinear process model selection
  (a) inputs: nonlinear deterministic models in CT, $f(x,u)$, and $h(x_k)$;
  (b) computation: integration of $f(x,u)$ from t_k to t_{k+1} to obtain $F(x_k,u_k)$;
  (c) output: nonlinear deterministic model in DT

$$x_{k+1} = F(x_k, u_k)$$
$$y_k = h(x_k)$$

(ii) stochastic disturbance model selection
  (a) inputs: integrating disturbance locations (process inputs, outputs, or combination of both);
  (b) computations: augment state vector with integrating disturbances as $\bar{x} = [x\ d]'$; integrate augmented model as in step (i) to obtain $F(\bar{x}_k, u_k)$, $G(\bar{x}_k)$, and $h(\bar{x}_k)$;
  (c) output: augmented nonlinear stochastic model in DT

$$\bar{x}_{k+1} = F(\bar{x}_k, u_k) + G(\bar{x}_k)w_k$$
$$y_k = h(x_k)$$

(iii) covariance identification from operating data
  (a) inputs: nonlinear stochastic model from step (ii), process operating data, ALS design parameters (N, N_d) , and set of initial guesses for covariances (Q,R);
  (b) computations: formulate and solve ALS problem using regular or diagonal technique;
  (c) output: covariance estimates (Q,R);

(iv) estimator selection and implementation
  (a) inputs: state estimators, e.g., MHE, EKF, among others (see Section Introduction and Prior work);
  (b) computations: select estimator type according to application and implement it;
  (c) output: implemented state estimator with noise statistics defined by ALS.

Conclusions

This article introduced a design method for nonlinear state estimation including the following steps: nonlinear stochastic modeling, noise covariance estimation, and estimator selection and implementation. Regarding the modeling approach, we showed that CT nonlinear SDEs can be accurately represented, for one step ahead predictions, by DT models from first principles with an added stochastic component that can be estimated from process operating data. On the covariance estimation, the LTV-ALS technique accurately estimated the process and measurement noise covariances for the addressed nonlinear examples (CSTR and ethylene copolymerization). This accuracy is justified by the fact that, in state estimation, EKF may not work well due to linearizations. In covariance estimation, however, linearization errors are indistinguishable from other error sources due to the central limit theorem, and thus estimated covariances are accurate enough to be used in state estimation. These covariances were used to systematically specify the noise statistics of the EKF and MHE state estimators. This provides an alternative for industrial practitioners to the current practice of determining KF and EKF gains arbitrarily by tuning, i.e., trial and error simulations, such that it gives satisfactory closed-loop performance. As suggested by the results in subsection Polymerization Process, physical models may be overly complex considering the available measurements; they may contain many unobservable and weakly observable modes. Overly complex structures lead to ill-conditioned or singular ALS problems for disturbance variance estimation. Ill-conditioning leads to unrealistic data demands for reliable estimates. As a solution to this problem, the design of a reduced-order extended Kalman filter to estimate only the strongly observable system states is under investigation. In general, overly complex disturbance models for weakly observable system models must be avoided. As future work, we plan to apply the steps described in this article to build and validate nonlinear state estimators for industrial chemical processes. Also, for such processes, we intend to quantify how often the covariance matrices change in an industrial data set by testing the time invariance of the disturbance statistics in the set as done previously in the literature with data from an industrial gas-phase reactor.[2] Finally, as a broader impact of this research, improving nonlinear state estimation in process monitoring and control will result in better chemical process operations giving rise to significant economic benefits.

Acknowledgments

The authors thank Dr. T. A. Soderstrom (ExxonMobil Chemical Company) for providing the polymerization model. The authors acknowledge the financial support from NSF through Grant CNS-0540147 and

from PRF through Grant 43321-AC9 and the support of ExxonMobil Chemical Company through the Texas-Wisconsin-California Control Consortium (TWCCC).

Literature Cited

1. Odelson BJ, Rajamani MR, Rawlings JB. A new autocovariance least-squares method for estimating noise covariances. *Automatica*. 2006;42:303–308.
2. Odelson BJ, Lutz A, Rawlings JB. The autocovariance least-squares method for estimating covariances: application to model-based control of chemical reactors. *IEEE Control Syst Technol*. 2006;14:532–541.
3. Rajamani MR, Rawlings JB. *Estimation of the Disturbance Structure from Data Using Semidefinite Programming and Optimal Weighting*. Technical Report 2007–02, TWMCC, Department of Chemical and Biological Engineering, University of Wisconsin–Madison, Madison, WI, 2007. Available at: http://jbrwww.che.wisc.edu/tech-reports.html.
4. Rajamani MR, Rawlings JB. Estimation of the disturbance structure from data using semidefinite programming and optimal weighting. *Automatica*. 2009;45:142–148.
5. Rajamani MR, Rawlings JB, Soderstrom TA. *Application of a New Data-Based Covariance Estimation Technique to a Nonlinear Industrial Blending Drum*. Technical Report 2007–03, TWMCC, Department of Chemical and Biological Engineering, University of Wisconsin–Madison, Madison, WI, 2007. Available at: http://jbrwww.che.wisc.edu/tech-reports.html.
6. Mehra RK. On the identification of variances and adaptive Kalman filtering. *IEEE Trans Automat Contr*. 1970;15:175–184.
7. Mehra RK. Approaches to adaptive filtering. *IEEE Trans Automat Contr*. 1972;17:903–908.
8. Carew B, Bélanger PR. Identification of optimum filter steady-state gain for systems with unknown noise covariances. *IEEE Trans Automat Contr*. 1973;18:582–587.
9. Bélanger PR. Estimation of noise covariance matrices for a linear time-varying stochastic process. *Automatica*. 1974;10:267–275.
10. Neethling C, Young P. Comments on Identification of optimum filter steady-state gain for systems with unknown noise covariances. *IEEE Trans Automat Contr*. 1974;19:623–625.
11. Noriega G, Pasupathy S. Adaptive estimation of noise covariance matrices in real-time preprocessing of geophysical data. *IEEE Trans Geosci Remote Sens*. 1997;35:1146–1159.
12. Åkesson BM, Jørgensen JB, Poulsen NK, Jørgensen SB. A generalized autocovariance least-squares method for Kalman filter tuning. *J Process Control*. 2008;18:769–779.
13. Kristensen NR, Madsen H, Jørgensen SB. Parameter estimation in stochastic grey-box models. *Automatica*. 2004;40:225–237.
14. Kristensen NR, Madsen H, Jørgensen SB. A method for systematic improvement of stochastic grey-box models. *Comput Chem Eng*. 2004;28:1431–1449.
15. Grancharova A, Kocijan J, Johansen TA. Explicit stochastic predictive control of combustion plants based on Gaussian process models. *Automatica*. 2008;44:1621–1631.
16. Kocijan J, Murray-Smith R. Nonlinear predictive control with a Gaussian process model. In: Murray-Smith R, Shorten R, editors. *Switching and Learning in Feedback Systems*, Volume 3355 of *Lecture Notes in Computer Science*. Heidelberg, Germany: Springer-Verlag, 2005:185–200.
17. Kocijan J, Likar B. Gas-liquid separator modelling and simulation with Gaussian process models. *Simulat Model Pract Theor*. 2008;16:910–922.
18. Rajamani MR. Data-Based Techniques to Improve State Estimation in Model Predictive Control. PhD Thesis, University of Wisconsin–Madison, Madison, WI, 2007.
19. Valappil J, Georgakis C. Systematic estimation of state noise statistics for extended Kalman filters. *AIChE J*. 2000;46:292–308.
20. Jazwinski AH. *Stochastic Processes and Filtering Theory*. New York: Academic Press, 1970.
21. Julier SJ, Uhlmann JK. Unscented filtering and nonlinear estimation. *Proc IEEE*. 2004;92:401–422.
22. Vachhani P, Narasimhan S, Rengaswamy R. Robust and reliable estimation via unscented recursive nonlinear dynamic data reconciliation. *J Process Control*. 2006;16:1075–1086.
23. Qu CC, Hahn J. Computation of arrival cost for moving horizon estimation via unscent Kalman filtering. *J Process Contr*. 2009;19:358–363.
24. Kandepu R, Imsland L, Foss BA. Constrained state estimation using the unscented kalman filter. In: *Procedings of the 16th Mediterranean Conference on Control and Automation*, Ajaccio, France, June, 2008:1453–1458.
25. Teixeira BOS, Tôrres LAB, Aguirre LA, Bernstein DS. Unscented filtering for interval-constrained nonlinear systems. In: *Proceedings of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, December 9–11, 2008:5116–5121.
26. Handschin JE, Mayne DQ. Monte Carlo techniques to estimate the conditional expectation in multistage nonlinear filtering. *Int J Control*. 1969;9:547–559.
27. Gordon NJ, Salmond DJ, Smith AFM. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proc F-Radar Signal Process*. 1993;140:107–113.
28. Doucet A, Godsill S, Andrieu C. On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat Comput*. 2000;10:197–208.
29. Arulampalam MS, Maskell S, Gordon NJ, Clapp T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans Signal Process*. 2002;50:174–188.
30. Chen WS, Bakshi BR, Goel PK, Ungarala S. Bayesian estimation of unconstrained nonlinear dynamic systems via sequential Monte Carlo sampling. *Ind Eng Chem Res*. 2004;43:4012–4025.
31. Lang L, Chen WS, Bakshi BR, Goel PK, Ungarala S. Bayesian estimation via sequential Monte Carlo sampling—constrained dynamic systems. *Automatica*. 2007;43:1615–1622.
32. Ungarala S, Li K, Chen Z. Constrained Bayesian state estimation using a cell filter. *Ind Eng Chem Res*. 2008;47:7312–7322.
33. Ungarala S, Li K. The use of a cell filter for state estimation in closed-loop NMPC of low dimensional systems. *J Process Control*. 2009;19:550–556.
34. Rao CV, Rawlings JB. Constrained process monitoring: moving-horizon approach. *AIChE J*. 2002;48:97–109.
35. Robertson DG, Lee JH, Rawlings JB. A moving horizon-based approach for least-squares state estimation. *AIChE J*. 1996;42:2209–2224.
36. Robertson DG, Lee JH. On the use of constraints in least squares estimation and control. *Automatica*. 2002;38:1113–1124.
37. Rao CV, Rawlings JB, Mayne DQ. Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations. *IEEE Trans Automat Contr*. 2003;48:246–258.
38. Alessandri A, Baglietto M, Battistelli G. Moving-horizon state estimation for nonlinear discrete-time systems: new stability results and approximation schemes. *Automatica*. 2008;44:1753–1765.
39. Rawlings JB, Bakshi BR. Particle filtering and moving horizon estimation. *Comput Chem Eng*. 2006;30:1529–1541.
40. Rajamani MR, Rawlings JB. Improved state estimation using a combination of moving horizon estimator and particle filters. In: *Proceedings of the American Control Conference*, New York, USA, July 11–13, 2007.
41. Zavala VM, Laird CD, Biegler LT. A fast computational framework for large-scale moving horizon estimation. In: DYCOPS, Vol. 3, Cancun, Mexico, June, 2007:21–30.
42. Zavala VM, Laird CD, Biegler LT. A fast moving horizon estimation algorithm based on nonlinear programming sensitivity. *J Process Contr*. 2008;18:876–884.
43. Kraus T, Kühl P, Wirsching L, Bock HG, Diehl M. A moving horizon state estimation algorithm applied to the Tennessee Eastman benchmark process. In: *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Heidelberg, Germany, 2006:377–382.
44. Diehl M, Ferreau HJ, Haverbeke N. Efficient numerical methods for nonlinear MPC and moving horizon estimation. In: *Proceedings of the International Workshop on Assessment and Future Directions of NMPC*, Pavia, Italy, September, 2008.
45. Haseltine EL, Rawlings JB. Critical evaluation of extended Kalman filtering and moving horizon estimation. *Ind Eng Chem Res*. 2005;44:2451–2460.
46. McAuley KB, MacGregor AE, Hamielec AE. A kinetic model for industrial gas-phase ethylene copolymerization. *AIChE J*. 1990;36:837–850.

47. Dadebo SA, Bell ML, McLellan PJ, McAuley KB. Temperature control of industrial gas phase polyethylene reactors. *J Process Contr*. 1997;7:83–95.
48. Gani A, Mhaskar P, Christofides PD. Fault-tolerant control of a polyethylene reactor. *J Process Control*. 2007;17:439–451.
49. Kloeden PE, Platen E. *Numerical Solution of Stochastic Differential Equations*. Berlin: Springer Verlag, 1992.
50. Jenkins GM, Watts DG. *Spectral Analysis and Its Applications*. 500 Sansome Street, San Fransisco, California: Holden-Day, 1968.
51. Sontag ED. *Mathematical Control Theory*, 2nd ed. New York: Springer-Verlag, 1998.
52. Schmidt SF. Application of state-space methods to navigation problems. In: Leondes CT editor. *Advances in Control Systems*, Vol. 3. New York: Academic Press, 1966:293–340.
53. Brown RG, Hwang PYC. *Introduction to Random Signals and Applied Kalman Filtering*. New York, Wiley, 1997.
54. Farrell JA, Barth M. *The Global Positioning System and Inertial Navigation*. McGraw-Hill, 1998.
55. Farrell JA. *Aided Navigation: GPS with High Rate Sensors*. McGraw-Hill, 2008.

Manuscript received Dec. 4, 2009, and revision received Apr. 27, 2010.