Covariance and State Estimation of Weakly Observable Systems: Application to Polymerization Processes

Fernando V. Lima, Murali R. Rajamani, Tyler A. Soderstrom, and James B. Rawlings, *Fellow, IEEE*

***Abstract*— Physical models for polymerization may be overly complex considering the available measurements, and they may contain many unobservable and weakly observable modes. Overly complex structures lead to ill-conditioned or singular problems for disturbance variance estimation. Ill conditioning leads to unrealistic data demands for reliable covariance estimates and state estimates. The goal of this paper is to build nonlinear state estimators for weakly observable systems, with focus on polymerization processes. State estimation requires knowledge about the noise statistics affecting the states and the measurements. These noise statistics are usually unknown and need to be estimated from operating data. We introduce a linear time-varying autocovariance least-squares (LTV-ALS) technique to estimate the noise covariances for nonlinear systems using autocorrelations of the data at different time lags. To reduce or eliminate the ill-conditioning problem, we design a reduced-order extended Kalman filter (EKF) to estimate only the strongly observable system states. This reduced filter, which is based on the Schmidt–Kalman filter, is used to perform the estimation of noise covariances by the LTV-ALS technique. Results of the implementation of the proposed method on a large-dimensional ethylene copolymerization example show that better conditioned state and covariance estimation problems can be obtained. We also show that high-quality state estimates can be obtained after the specification of the noise statistics of EKF estimators by ALS.**

***Index Terms*— Autocovariance least squares (ALS), nonlinear state estimation, nonlinear stochastic modeling, time-varying systems, weakly observable systems.**

I. Introduction

ADVANCED feedback control schemes for nonlinear models, such as model predictive control, often have a state estimator that reconciles past inputs and plant measurements to make an estimate of the current state of the system. The regulator then uses the current state estimate and the model to optimize future control inputs. Thus, the performance of an advanced feedback control system is directly affected by the quality of the state estimates. For optimal performance, the state estimator requires knowledge of the noise statistics affecting the plant. If these noises are modeled as zero-mean Gaussian sequences, then the covariances are required to specify their statistics. In practice, these covariances are typically unknown and therefore chosen arbitrarily by process control engineers to get satisfactory closed-loop performance. Most chemical processes are characterized by both nonlinearity in the dynamics and significant levels of process and sensor noise. Also, most complex product properties are not measurable but must be inferred from other measurements combined with nonlinear property models. Specifically, polymerization process models, in addition to containing many unobservable and weakly observable modes, are nonlinear and large dimensional (around 50 states and 20 measurements), the product properties of most interest are complex nonlinear functions of the process state, and the stochastic disturbance structure is unknown *a priori*. The objective of this paper presented here is to design state estimators for weakly observable systems with particular interest in polymerization process examples. A gas-phase ethylene copolymerization process model from the literature [1]–[3] is addressed here. This task is challenging, as it requires the following steps: nonlinear process model selection, stochastic disturbance model selection, covariance identification from operating data, determination of the subset of the state to be estimated online from the measurements, and estimator selection and implementation.

Regarding the modeling task, we use fully nonlinear stochastic models in discrete time obtained by 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 stochastic component estimated from a routine set of operating data that provide a typical sample of the measurement and process disturbances affecting the system. Lima and Rawlings [4] have shown that a continuous-time nonlinear stochastic model for the states [e.g., a stochastic differential equation (SDE) system model] can be well represented by this discrete-time (DT) model structure. Integrating disturbance models are used to provide offset-free control of the properties of interest

Manuscript received February 27, 2011; revised April 6, 2012; accepted April 14, 2012. Manuscript received in final form May 14, 2012. Date of publication June 19, 2012; date of current version June 14, 2013. This work was supported by the NSF under Grant CNS-0540147, the PRF under Grant 43321-AC9, and the ExxonMobil Chemical Company through the Texas-Wisconsin-California Control Consortium. Recommended by Associate Editor A. Alessandri.

F. V. Lima and J. B. Rawlings are with the Department of Chemical and Biological Engineering, University of Wisconsin, Madison, WI 53706 USA (e-mail: fvlima@umn.edu; rawlings@engr.wisc.edu).

M. R. Rajamani was with BP Research and Technology, Naperville, IL 60563 USA. He is now with the Booth School of Business, University of Chicago, Chicago, IL 60637 USA (e-mail: murali.rajamani@bp.com).

T. A. Soderstrom is with ExxonMobil Chemical Company, Baytown Technology and Engineering Complex, Baytown, TX 77522 USA (e-mail: tyler.a.soderstrom@exxonmobil.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCST.2012.2200296

1063-6536/$31.00 © 2012 IEEE

while maintaining a low enough complexity so that the disturbance statistics can be determined from the available measurements. For the estimation of these statistics, specifically the covariances of the process (Q) and measurement (R) noises, [5] and [6] proposed the autocovariance least-squares (ALS) technique for linear models. Here, we extend this technique to address nonlinear and time-varying models. The proposed linear time-varying autocovariance least-squares (LTV-ALS) technique uses routine process operating data and thus does not require input–output testing to be applied to the system. Simply stated, the general idea of this technique is that the state noise w_k gets propagated in time, but the measurement noise v_k applies only at sampling times and thus is not propagated in time. Hence, taking autocovariances of the data at different time lags separately gives the covariances of w_k (Q) and v_k (R) (see Section II-C for the LTV-ALS technique formulated for nonlinear models). Refer to [4] for more information on other covariance estimation approaches from the literature for linear and nonlinear systems. The estimated noise covariances are used to systematically specify the noise statistics of the selected nonlinear state estimator, e.g., the extended Kalman filter (EKF) [7] or the moving horizon estimator [8]–[12] (see [13] and [4] for an overview on many other nonlinear state estimation methods).

In general, overly complex disturbance models for weakly observable system models must be avoided. To reduce or eliminate this ill-conditioning problem that may also plague the state estimation step, we design a reduced-order EKF to estimate only the strongly observable system states. This reduced filter is used to perform the ALS estimation of noise covariances. One example of such a filter is the Schmidt–Kalman filter (SKF), which was originally developed for navigation systems to improve numerical stability and reduce the computational complexity of Kalman filters [14], [15], and later used to tackle weakly observable systems [16], [17]. The general idea of this technique is to remove weakly observable states in the Kalman filter gain calculation, producing a filter that does not estimate the removed state variables but still keeps track of the influences these states have on the gain applied to the other states.

The outline of the rest of this paper is as follows. First, a summary of the steps necessary to design nonlinear state estimators is presented. This summary includes the proposed LTV-ALS technique. Then, a state estimator is designed for the polymerization process mentioned above and the difficulties encountered during the covariance and state estimation steps are reported. To overcome these difficulties, a novel filtering approach for weakly observable systems based on the SKF is introduced. Next, the reduced filter is designed and implemented in the same process to show that better conditioned state and covariance estimation problems can be obtained. We also show that high-quality state estimates can be obtained after the specification of the noise statistics of EKF estimators by ALS. Finally, conclusions are presented.

II. Design Steps for Nonlinear State Estimators

Here we provide some background on the following steps necessary to design nonlinear state estimators for weakly observable systems: stochastic modeling and covariance identification from operating data.

A. Nonlinear Stochastic Modeling

As briefly mentioned above, the modeling approach used here relies on the combination of a nonlinear deterministic model obtained from the integration of a first-principles model and a stochastic component estimated from process operating data that is discrete, available at every process sampling time $\Delta_k = t_{k+1} - t_k$, and contains the process and measurement noise [4].

The deterministic part consists of the following nonlinear model in DT:

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

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, $F(x_k, u_k)$ is obtained by integrating the deterministic nonlinear model $f(x, u)$ from t_k to t_{k+1} using an ordinary differential equation solver with a zero-order hold on the input u_k.

Regarding the stochastic part of this model, first the state vector is augmented with an integrated white noise component $d \in \mathbb{R}^q$, as $\tilde{x} = [x \; d]'$. This noise component is added to cope with potential plant–model mismatches and to achieve offset-free performance for the outputs [18]–[20]. Specifically, the number of integrating disturbances added as well as their location (process inputs, outputs or combination of both) are user dependent [21]. To ensure zero offset in all the process outputs, a general recommendation is to add a number of integrated disturbances that correspond to the number of outputs/measurements ($q = p$) [20]. The deterministic part of the augmented model, obtained as (1) using $\tilde{x}$, with an added noise term, results in the following nonlinear stochastic model in discrete time:

$$\tilde{x}_{k+1} = \tilde{F}(\tilde{x}_k, u_k) + \tilde{G}(\tilde{x}_k) w_k \qquad y_k = h(\tilde{x}_k) + v_k \tag{2}$$

in which $w \in \mathbb{R}^g$ are the process noises and $v \in \mathbb{R}^p$ are the measurement noises, which are assumed to be Gaussian with mean zero and time-invariant covariance matrices $w_k \sim N(0, Q)$ and $v_k \sim N(0, R)$, respectively.

Lima and Rawlings [4] have shown that, for one-step-ahead predictions, this nonlinear stochastic model structure is accurate for state estimation and feedback control, even if a SDE model [22] is the plant generating the data. They also showed, through a reactor case study, that the noises used in such a model structure are well approximated by normal distributions with time invariant statistics. Thus, in the next sections, (2) will be used for covariance and state estimation. Also, for presentation purposes, we will drop the symbol $\tilde{}$ from $\tilde{x}$, $\tilde{F}$, and $\tilde{G}$.

B. Noise Covariance Estimation for Nonlinear Systems

With the chosen nonlinear process model and stochastic disturbance model structures, we next turn to the task of estimating the process and output noise variances given only routine process output operating data. In this and the next

subsections, we present the LTV-ALS technique for nonlinear systems.

Recall the DT nonlinear stochastic model structure (2)

$$x_{k+1} = F(x_k, u_k) + G(x_k)w_k \qquad y_k = h(x_k) + v_k \tag{3}$$

in which $w_k \sim N(0, Q)$ and $v_k \sim N(0, R)$ are uncorrelated with each other.

Let the state estimates be obtained starting from an arbitrary initial value using a time-varying stable filter gain sequence L_k. An example of such a filter gain would be the sequence obtained by implementing the EKF. The only condition on the time-varying gains L_k is that they are stable (see Assumption 1). The state estimation can then be described by the following equations:

$$\begin{aligned} \hat{x}_{k+1|k} &= F(\hat{x}_{k|k}, u_k) \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + L_k(y_k - \hat{y}_{k|k-1}) \\ \hat{y}_{k|k-1} &= h(\hat{x}_{k|k-1}) \end{aligned} \tag{4}$$

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 (3) around $(\hat{x}_{k|k}, u_k)$ is performed, then this model can be represented by the following set of time-varying equations:

$$x_{k+1} = A_k x_k + B_k u_k + G_k w_k \qquad y_k = C_k x_k + v_k \tag{5}$$

in which

$$\begin{aligned} A_k &= \left.\frac{\partial F(x_k, u_k)}{\partial x_k}\right|_{(\hat{x}_{k|k}, u_k)}, \\ C_k &= \left.\frac{\partial h(x_k)}{\partial x_k}\right|_{\hat{x}_{k|k}} \\ G_k &= G(\hat{x}_{k|k}, u_k). \end{aligned}$$

Also, the linearized estimation equations are given by

$$\begin{aligned} \hat{x}_{k+1|k} &= A_k \hat{x}_{k|k} + B_k u_k \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + L_k(y_k - \hat{y}_{k|k-1}) \\ \hat{y}_{k|k-1} &= C_k \hat{x}_{k|k-1}. \end{aligned} \tag{6}$$

Subtracting the predicted state estimate and output in (6) from the linearized model in (5), we obtain the following approximate time-varying linear model for the innovations:

$$\begin{aligned} \varepsilon_{k+1} &\approx \underbrace{(A_k - A_k L_k C_k)}_{\bar{A}_k}\varepsilon_k + \underbrace{\begin{bmatrix} G_k & -A_k L_k \end{bmatrix}}_{\bar{G}_k}\underbrace{\begin{bmatrix} w_k \\ v_k \end{bmatrix}}_{\bar{w}_k} \\ \mathcal{Y}_k &\approx C_k \varepsilon_k + v_k \end{aligned} \tag{7}$$

in which $\varepsilon_k = (x_k - \hat{x}_{k|k-1})$ denotes the state estimate error and $\mathcal{Y}_k = (y_k - \hat{y}_{k|k-1})$ the innovations sequence at time t_k. The noises w_k, v_k driving the innovations sequence are assumed to be drawn from time-invariant covariances Q, R. The next subsection presents a technique for estimating the covariances Q, R using autocovariances of data at different time lags.

C. Time-Varying ALS Technique

The ALS covariance estimation technique described in [5] and [6] was applied to a linear time-invariant model. When using nonlinear or time-varying models, a key difference is that the estimate error covariance $P_k = E[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})]$ is the time-varying solution to the Riccati equation and does not reach a steady-state value. No simple equation can then be written for P_k in terms of Q, R and the system matrices as in the linear time-invariant case. The following Assumption 1, however, allows the extension of the ALS technique to time-varying and nonlinear systems.

In the assumption below, and in the rest of this paper, we define for $i \geq j$

$$\bar{A}_{[i,j]} = \bar{A}_i \bar{A}_{i-1} \cdots \bar{A}_{j+1} \bar{A}_j$$

and $\bar{A}_{[i,j]} = I$ for $i < j$.

Assumption 1: The time-varying filter gain sequence L_k used in (4) is such that, when used in the approximate linearization given by (7), it produces a sequence of $\bar{A}_k = (A_k - A_k L_k C_k)$ matrices such that the product $\bar{A}_{[k-1,0]} \to 0$ as k increases. This assumption can be guaranteed under suitable detectability and stabilizability conditions [23] and is necessary to fulfill the steady-state assumption of ALS [5], [6].

Starting from an arbitrary initial condition ε_0 at t_0, consider the evolution of (7) up to time t_k to obtain the following:

$$\begin{aligned} \mathcal{Y}_k &= C_k\left(\bar{A}_{[k-1,0]}\right)\varepsilon_0 + C_k\left(\bar{A}_{[k-1,1]}\bar{G}_0\bar{w}_0 \right. \\ &\quad \left. + \bar{A}_{[k-1,2]}\bar{G}_1\bar{w}_1 + \cdots + \bar{G}_{k-1}\bar{w}_{k-1}\right) + v_k \\ \mathcal{Y}_{k+j} &= C_{k+j}\left(\bar{A}_{[k+j-1,0]}\right)\varepsilon_0 + C_{k+j}\left(\bar{A}_{[k+j-1,1]}\bar{G}_0\bar{w}_0 \right. \\ &\quad \left. + \bar{A}_{[k+j-1,2]}\bar{G}_1\bar{w}_1 + \cdots + \bar{G}_{k+j-1}\bar{w}_{k+j-1}\right) + v_{k+j}. \end{aligned} \tag{8}$$

The covariance of $\bar{w}_k = \left[\begin{smallmatrix} w_k \\ v_k \end{smallmatrix}\right]$ is given by $\mathbb{Q} = \left[\begin{smallmatrix} Q & 0 \\ 0 & R \end{smallmatrix}\right]$.

From Assumption 1, the index $k > 0$ is chosen to be large enough such that $\bar{A}_{[k-1,0]} \approx 0$. The effect of the initial estimate error ε_0 is then negligible in $\mathcal{Y}_{k+j}$ for $j > 0$. The terms involving ε_0 in (8) can then be neglected. From (8), we obtain the following expressions for the expectation of the autocovariances at different lags:

$$\begin{aligned} E[\mathcal{Y}_k\mathcal{Y}_k'] &= C_k\left[\bar{A}_{[k-1,1]}\bar{G}_0 \cdots \bar{G}_{k-1}\right] \\ &\quad \times \bigoplus_{i=1}^{k}\mathbb{Q}\left[\bar{A}_{[k-1,1]}\bar{G}_0 \cdots \bar{G}_{k-1}\right]' C_k' + R \\ E[\mathcal{Y}_{k+j}\mathcal{Y}_k'] &= C_{k+j}\left[\bar{A}_{[k+j-1,1]}\bar{G}_0 \cdots \bar{A}_{[k+j-1,k]}\bar{G}_{k-1}\right] \\ &\quad \times \bigoplus_{i=1}^{k}\mathbb{Q}\left[\bar{A}_{[k-1,1]}\bar{G}_0 \cdots \bar{G}_{k-1}\right]' C_k' \\ &\quad - C_{k+j}\left(\bar{A}_{[k+j-1,k+1]}\right) A_k L_k R. \end{aligned} \tag{9}$$

In the above equation and in the remainder of this paper, we use the standard properties and symbols of Kronecker products and direct sum [24], [25]. Specifically, the symbol $\otimes$ is the standard symbol for the Kronecker product and the symbol $\bigoplus$ represents the direct sum that satisfies the following property:

$$\bigoplus_{i=1}^{k}\mathbb{Q} = (I_k \otimes \mathbb{Q}).$$

Remark 1: In general, instead of starting from the initial condition ε_0 at t_0, we can start from an initial condition ε_m at t_m and calculate the expectations as above, provided the indices k and m are such that the following condition holds as in Assumption 1:

$$\bar{A}_{[m+k-1,m]} \approx 0.$$

Let the autocovariance matrix $\mathscr{R}_k(N)$ be defined as the expectation of the innovations data at different time lags over a user-defined window N [26]

$$\mathscr{R}_k(N) = E\begin{bmatrix} \mathscr{Y}_k\mathscr{Y}_k' \\ \vdots \\ \mathscr{Y}_{k+N-1}\mathscr{Y}_k' \end{bmatrix}. \tag{10}$$

Using (8) and (10), we obtain

$$\mathscr{R}_k(N) = \underbrace{\begin{bmatrix} I_p \\ -C_{k+1}A_kL_k \\ \vdots \\ -C_{k+N-1}\left(\bar{A}_{[k+N-2,k+1]}\right)A_kL_k \end{bmatrix}}_{\Psi_1} R$$
$$+\Gamma\Omega_1\bigoplus_{i=1}^{k} Q\Omega_1'\Gamma_1' + \Gamma\Omega_2\bigoplus_{i=1}^{k} R\Omega_2'\Gamma_1' \tag{11}$$

in which the matrices are defined as follows and dimensioned appropriately:

$$\Gamma = \begin{bmatrix} C_k\left(\bar{A}_{[k-1,1]}\right) & \cdots & C_k \\ \vdots & \ddots & \vdots \\ C_{k+N-1}\left(\bar{A}_{[k+N-2,1]}\right) & \cdots & C_{k+N-1}\left(\bar{A}_{[k+N-2,k]}\right) \end{bmatrix} \tag{12}$$

and

$$\Omega_1 = \begin{bmatrix} G_0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & G_{k-1} \end{bmatrix}, \tag{13}$$

$$\Omega_2 = \begin{bmatrix} -A_0L_0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & -A_{k-1}L_{k-1} \end{bmatrix}$$

$$\Gamma_1' = \begin{bmatrix} \left(\bar{A}_{[k-1,1]}\right)'C_k' \\ \vdots \\ C_k' \end{bmatrix}. \tag{14}$$

Also, Γ_1 is the first row block of the Γ matrix.

We use the subscript "s" to denote the columnwise stacking of the elements of a matrix into a vector. Stacking both sides of (11), we obtain

$$[\mathscr{R}_k(N)]_s = (\Gamma_1\Omega_1 \otimes \Gamma\Omega_1)\mathcal{I}_{g,k}(Q)_s$$
$$+\left[(\Gamma_1\Omega_2 \otimes \Gamma\Omega_2)\mathcal{I}_{p,k} + I_p \otimes \Psi_1\right](R)_s. \tag{15}$$

Here, $\mathcal{I}_{p,k}\epsilon\mathbb{R}^{(pk)^2\times p^2}$ is a permutation matrix to convert the direct sum to a vector. This matrix contains 1s and 0s and satisfies the following relation:

$$\left(\bigoplus_{i=1}^{k} R\right)_s = \mathcal{I}_{p,k}(R)_s.$$

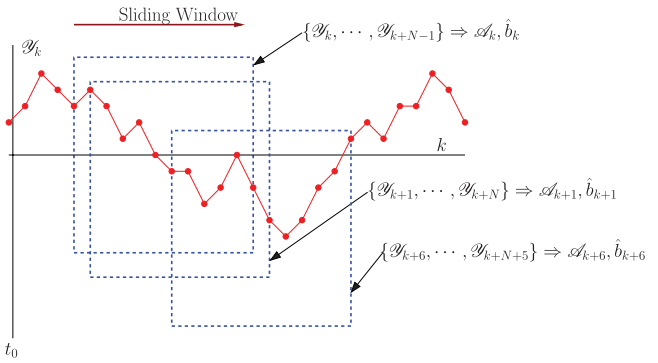


Fig. 1. Strategy for calculating the time-varying quantities $\mathscr{A}_k$ and $\hat{b}_k$ in (17) and (18).

If we have an estimate of the autocovariance matrix $\mathscr{R}_k(N)$, denoted by $\hat{\mathscr{R}}_k(N)$, and let $\hat{b}_k = [\hat{\mathscr{R}}_k(N)]_s$, then from (15) we can formulate a positive semidefinite constrained least-squares problem in the unknown covariances Q, R [5]. The optimization to be solved is given by

$$\min_{Q,\,R} \left\| \mathscr{A}_k \begin{bmatrix} (Q)_s \\ (R)_s \end{bmatrix} - \hat{b}_k \right\|_W^2$$
$$\text{s.t.} \quad Q, R \geq 0, \quad Q = Q', R = R' \tag{16}$$

in which

$$\mathscr{A}_k = \begin{bmatrix} \mathscr{A}_{k1} & \mathscr{A}_{k2} \end{bmatrix}$$
$$\mathscr{A}_{k1} = (\Gamma_1\Omega_1 \otimes \Gamma\Omega_1)\mathcal{I}_{g,k}$$
$$\mathscr{A}_{k2} = \left[(\Gamma_1\Omega_2 \otimes \Gamma\Omega_2)\mathcal{I}_{p,k} + I_p \otimes \Psi_1\right]. \tag{17}$$

We will refer to the optimization in (16) as the ALS technique. The necessary and sufficient conditions for the uniqueness of the ALS optimization in (16) are given in [6]. For the estimates of Q, R in the ALS optimization to have minimum variance, the weighting matrix W in the ALS objective is given by $W = T_k^{-1}$, in which T_k is the covariance of $\hat{b}_k$ [27].

The matrices $\mathscr{A}_k$ and the vector $\hat{b}_k$ in (16) have the time subscript "k" to emphasize that these quantities are time-varying and based on the time-varying approximation given in (7).

As the only dataset available for estimating the time-varying quantity $\mathscr{R}_k(N)$ defined in (10) is $\{\mathscr{Y}_k, \ldots, \mathscr{Y}_{k+N-1}\}$, the only calculable estimate of $\hat{b}_k$ is given by

$$\hat{b}_k = [\hat{\mathscr{R}}_k(N)]_s = \begin{bmatrix} \mathscr{Y}_k\mathscr{Y}_k' \\ \vdots \\ \mathscr{Y}_{k+N-1}\mathscr{Y}_k' \end{bmatrix}_s. \tag{18}$$

At every time instant t_k, we compute the quantities $\mathscr{A}_k$ and $\hat{b}_k$ from (17) and (18). To simplify the computation, the matrices $\Gamma, \Omega_1, \Omega_2, \Gamma_1$ defined in (12) and (13) and used in the calculation of $\mathscr{A}_k$ can be computed starting from an initial condition at time t_m rather than t_0 as given in Remark 1. We then use a sliding-window strategy to compute the time-varying matrices $\mathscr{A}_k$ and $\hat{b}_k$. Fig. 1 illustrates the calculation procedure.

Using the computed time-varying matrices and the ALS formulation in (16), we can then solve the following

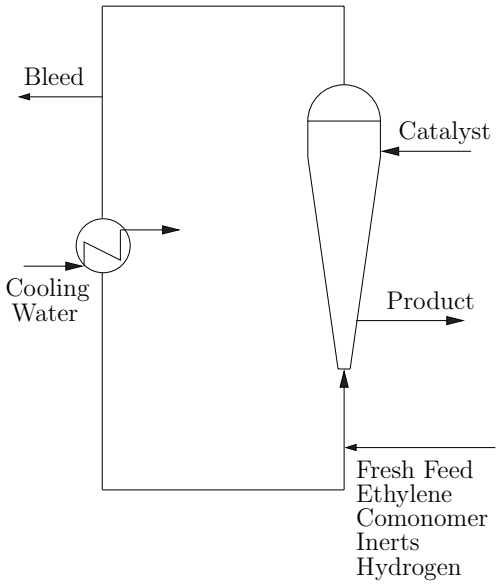


Fig. 2. Gas-phase ethylene copolymerization reactor system [2].

optimization for a set of data of length N_d to estimate Q, R:

$$\min_{Q,\,R} \left\| \begin{bmatrix} \mathscr{A}_k \\ \vdots \\ \mathscr{A}_{N_d-N+1} \end{bmatrix} \begin{bmatrix} (Q)_s \\ (R)_s \end{bmatrix} - \begin{bmatrix} \hat{b}_k \\ \vdots \\ \hat{b}_{N_d-N+1} \end{bmatrix} \right\|_{W_f}^2$$
$$\text{s.t.} \quad Q, R \geq 0, \quad Q = Q', R = R'. \tag{19}$$

Since $\hat{b}_k, \hat{b}_{k+1}, \ldots$ are not independent, the weighting matrix W_f is not block-diagonal. The formula for W_f is a complicated function of the unknown covariances Q, R and an iterative procedure is required for its calculation [6]. Thus, we use $W_f = I$ to avoid the computationally infeasible and nonlinear calculation. Finally, since no explicit solution to the semidefinite optimization problem in (19) exists, it is solved with a logarithmic barrier function to ensure that the calculated Q and R are positive semidefinite matrices [28]. Typical semidefinite programming algorithms have polynomial worst case complexity [29].

Remark 2: Notice that, if the system is time-invariant, we have $\mathscr{A}_k = \cdots = \mathscr{A}_{N_d-N+1}$ and we can then recover the time-invariant ALS optimization presented in [5] and [6]. Additionally, the linear time-invariant technique can be applied to the nonlinear case when the matrices $\mathscr{A}_k, \ldots, \mathscr{A}_{N_d-N+1}$ are equal or do not change significantly.

Remark 3: As the value of the ALS design parameter, i.e., the horizon length (N), increases, the accuracy of the covariance estimation improves at the expense of a larger computational cost to perform such an estimation. This parameter value is selected based on the application. For example, for the polymerization process presented in Sections III and IV, the value of $N = 15$ is large enough to provide good covariance estimates.

Remark 4: Assumption 1 is a simple practical requirement that the time-varying linear approximation of the full nonlinear model has a gain sequence that makes the estimate error asymptotically zero. This requirement is satisfied in most industrial applications that use a linear approximation to design the state estimator.

Remark 5: Note that this approach requires linearizations of the nonlinear model around the current state estimate and input values. This linearization is not a critical issue for noise covariance estimation because these linearization errors are indistinguishable from other error sources as a result of the central limit theorem. Thus, the estimated covariances are accurate enough to be used in state estimation [4].

Remark 6: 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 (16) are ill conditioned. This estimation is also useful to increase the speed of the ALS computations [28]. A mathematical summary of the diagonal ALS technique can be found in [4].

III. Results: Ethylene Copolymerization Process Example

In this section, we present the results of the application of the design steps for nonlinear state estimation described above to a gas-phase ethylene copolymerization process model [1]–[3]. This polymerization process has the following variables: 41 states, 17 measurements, 5 inputs, and 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). In addition to the typical challenges already mentioned regarding polymerization process models (such as large dimensionality, presence of nonlinearities, and weakly/unobservable modes), the presence of grade transition policies in this copolymerization process [3] also motivates the application of the developed time-varying covariance estimation technique. Fig. 2 shows a schematic diagram of this process (see [1]–[3] for additional information about this process).

In order to model this polymerization system, a nonlinear first-principles model with an added noise component was developed, as explained in Section II-A. This model was built on the basis of the cited references above and has the structure of (2). Specifically, the matrix G_k was computed on the basis of the sensitivities of the first-principles model equations with respect to changes in the process noise components added to this model. To generate the simulated data from the process model 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 $w_{\text{sim}} \sim (0, Q_{\text{sim}})$ 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 $w_k \sim (0, Q)$ and $v_k \sim N(0, R)$ for the model structure selected and

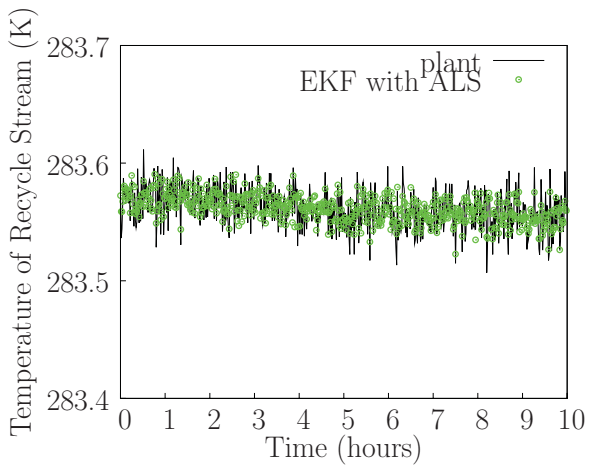


Fig. 3. EKF implementation results for polymerization problem using simulated data (indicated as "plant").

parameter values of $N = 15$ and $N_d = 6000$ are given by

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

Also, to satisfy the steady-state assumption of ALS (see Assumption 1), a value of $k = 200$ was selected for this case study. Note that the LTV-ALS technique estimates these covariances accurately. Fig. 3 shows the temperature of the recycle streams (in K) plots for the EKF with the statistics defined by ALS. The accuracy of the ALS estimates is also demonstrated by the low value of the calculated mean absolute relative error (MARE) between plant (simulated data) and state estimates for the variable in this figure (MARE $= 2.66 \times 10^{-5}$). However, the formulated full ALS problem is ill conditioned; thus, we could only estimate accurately the variance of a single process disturbance and the diagonal components of the measurement covariance matrix using the diagonal ALS technique mentioned above [4]. 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. 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 (16). An overly complex disturbance model for the available information in the measurements is detected by poor conditioning in the $\mathscr{A}_k$ matrices. In addition to this issue on covariance estimation, ill conditioning may also plague the state estimation step. Poor conditioning can be reduced or eliminated by rescaling the process model and designing a reduced-order Kalman filter to estimate only the strongly observable system states. This filter used for covariance estimation is designed next.

IV. Weakly Observable Systems

As briefly mentioned above, the general idea of the SKF technique is to remove weakly observable states in the Kalman filter gain calculation, producing a well-conditioned filter that does not estimate the removed state variables but still keeps track of the influences these states have on the gain applied to the other states. In this section, we describe the procedure to design a reduced filter for weakly observable systems and apply this technique to estimate the covariances for the ethylene copolymerization process. This performed estimation is well-conditioned because only strongly observable modes are involved. By weakly observable modes, we mean the modes of a linear system that are nearly unobservable. After a linear system is transformed into its observability canonical form, the weakly observable modes are those modes whose singular values of the observability matrix are positive, and hence observable, but less than some chosen small threshold.[1]

A. Filtering Approach: SKF

The SKF design procedure is as follows:

1) first, have the LTV system (5) in its observability canonical form

$$\left[\frac{x_o}{x_{\text{no}}}\right]_{k+1} = \underbrace{\left[\begin{array}{c|c} A_o & 0 \\ \hline A_{21} & A_{\text{no}} \end{array}\right]_k}_{A^c{}_k} \underbrace{\left[\frac{x_o}{x_{\text{no}}}\right]_k}_{x^c{}_k} + \underbrace{\left[\frac{B_o}{B_{\text{no}}}\right]_k}_{B^c{}_k} u_k + \underbrace{\left[\frac{G_o}{G_{\text{no}}}\right]_k}_{G^c{}_k} w_k$$

$$y_k = \underbrace{\left[C_o \; 0\right]_k}_{C^c{}_k} \left[\frac{x_o}{x_{\text{no}}}\right]_k + v_k$$

in which

$$x^c{}_k = T_k' x_k, \quad A^c{}_k = T_k' A_k T_k, \quad B^c{}_k = T_k' B_k,$$
$$G^c{}_k = T_k' G_k, \quad C^c{}_k = C_k T_k$$

and T_k is a similarity transformation matrix that satisfies the property $T_k^{-1} = T_k'$. The observability canonical form, including the similarity transformation matrix, is computed using the QR decomposition method [30]. In this structure, $x_o \in \mathbb{R}^o$ are the observable modes and $x_{\text{no}} \in \mathbb{R}^{n_o}$ are the unobservable/weakly observable modes. Also, at every k, the pair (A_o, C_o) is assumed observable and the system is still detectable if $\lambda(A_{\text{no}}) \leq 1$;

2) for every k, calculate the KF gain for the observable system, L_o, using A_o, C_o, and G_o. This gain is associated with the strongly observable modes;

3) augment L_o with n_o rows of zeros, corresponding to the unobservable/weakly observable modes as $L^c = \left[\frac{L_o}{O_{n_o \times p}}\right]$;

[1] Note that this meaning should not be confused with the concept of a weakly observable *nonlinear* system. A weakly observable nonlinear system is one in which the state x can be distinguished from other states in a full neighborhood of x [31, p. 733]. If the reader prefers to avoid this multiple usage, substitute "nearly unobservable" in place of our use of "weakly observable" in this paper.

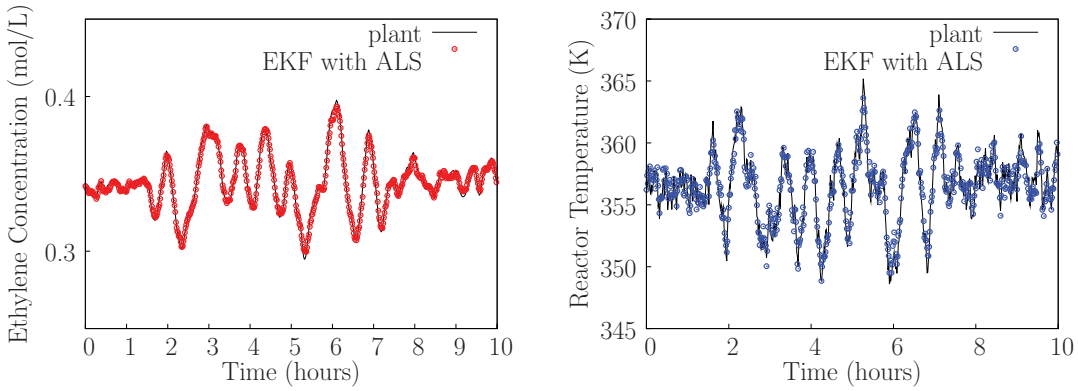


Fig. 4. SKF implementation results for polymerization problem using simulated data (indicated as "plant").

4) finally, the physical states are calculated using the new reduced filter as

$$\hat{x}_{k+1|k+1} = F(\hat{x}_{k|k}, u_k) + L_{k+1}\left[y_{k+1} - h(\hat{x}_{k+1|k})\right]$$

in which $L_{k+1} = T_k L^c$.

The estimator obtained is well conditioned even if the original system had several weakly observable and unobservable modes. As there are only observable modes involved in the KF gain calculation, the convergence properties of the obtained filter are similar to the classical KF [32].

Remark 7: Application of SKF procedure to linear time-invariant systems: Even though the SKF procedure presented here shows time-varying matrices, this procedure could be performed for the time-invariant case by simply using the corresponding time-invariant matrices. In either case, the calculated gains are used for covariance estimation by the ALS technique. Calculating the gain at each time step provides more accurate values at a higher computational cost. However, depending on how much the system matrices change with time, in some cases, only one KF gain may be accurate enough for covariance estimation.

B. SKF results: Ethylene Copolymerization Revisited

Here we present the covariance and state estimation results after rescaling the process model and subsequently applying the SKF approach described above to the ethylene copolymerization example introduced in Section III. Specifically, rescaling was performed by dividing each state and output variable by its steady-state value, which was necessary due to the stiff nature of the system of equations that describe the process model.

Consider the case in which one zero-mean Gaussian process noise was added to each of the following state variables: hydrogen concentration, reactor temperature, and the cooler cooling water temperature (three process noises in total). Also, sensor noises were added to all the measurements as in the previous case. For data generation, the following noise covariances were assumed:

$$Q_{\text{sim}} = \text{diag}(5 \times 10^{-11}, 2 \times 10^{-4}, 0.5)$$
$$R_{\text{sim}} = \text{diag}(0.02, 5 \times 10^{-4}, 6 \times 10^{4}, 0.3, 80, 0.07, 40, 1.5 \times 10^{-5}, 2 \times 10^{-4}, 5 \times 10^{-5}, 1 \times 10^{-6}, 1 \times 10^{-13}, 0.2, 6 \times 10^{-4}, 100, 50, 55).$$

Applying the ALS technique, the following estimated values are obtained:

$$Q_{\text{als}} = \text{diag}(5 \times 10^{-11}, 1.8 \times 10^{-4}, 0.43)$$
$$R_{\text{als}} = \text{diag}(0.03, 5.1 \times 10^{-4}, 6 \times 10^{4}, 0.3, 84, 0.07, 46, 1.9 \times 10^{-5}, 2.4 \times 10^{-4}, 5.9 \times 10^{-5}, 1 \times 10^{-6}, 1.5 \times 10^{-13}, 0.5, 6.1 \times 10^{-4}, 114, 60, 59).$$

Notice that, once again, the covariances were accurately estimated by ALS. Fig. 4 shows the plots for two important variables of this process, i.e., ethylene concentration and reactor temperature, in which the state estimates are obtained by implementing EKF with statistics determined by ALS. The accuracy of the ALS estimates is once again demonstrated by the low MARE values between plant and state estimates for both variables in this figure (5.19×10^{-3} and 1.59×10^{-3}, respectively).

Remark 8: To build the initial EKF gain sequence L_k and subsequently the LTV-ALS matrices, educated guesses for the covariance matrices are necessary. These guesses can be calculated using the variances of measured variables in representative sections of their datasets. For the specific case study above, 17 measured variables are present. For the three output variables that are directly associated with the three process noises added, such as the hydrogen composition in the reactor (calculated based on concentration measurements of all species), the reactor temperature, and the cooling water outlet temperature, half of the variance in the measured data, σ^2, was attributed to the R variance, and the other half to the Q variance (corrected by the model matrices at steady state that express the relationship between the output and the state noises). Thus, for such variables, the diagonal elements of the guessed covariance matrices (Q_g, R_g) were computed as follows:

$$R_g(i,i) = \frac{\sigma^2(i)}{2}, \quad Q_g(l,l) = \frac{\sigma^2(i)}{2} \times \frac{1}{[C_k(i,j)G_k(j,l)]^2}$$

in which the indexes i, j, and l correspond to the outputs, states with added noise and process noises, respectively.

The remaining 14 components of the diagonal of the R_g matrix were selected as the measurement variances associated with their corresponding outputs, $R(i,i) = \sigma^2(i)$. Also, the off-diagonal terms ($i \neq j$) of both matrices were set to 0.

There are several unobservable and weakly observable modes in this application, and their number depends on the tolerance selected for the QR decomposition method used to compute the observability canonical form. For example, for a tolerance of 10^{-2}, 25 modes were identified. Specifically, these modes correspond to states primarily associated with the Ziegler–Natta polymerization model, such as the polymer moments and number of sites, and were suppressed in the SKF gain calculation. Thus, after rescaling the process model and applying the SKF approach, better conditioned state and covariance estimation problems are obtained. Specifically, before performing these two steps, the condition number of the $\mathscr{A}_k$ matrices of the least-squares problem defined by (16) was 2.13×10^{19} for the ALS design parameter values of $N = 15$ and $N_d = 6000$. After the implementation of the proposed approach, this number is reduced to 83.85 for the same parameter values. Thus, the ALS technique is now able to estimate multiple process noises as opposed to only one noise as in the previous case. It is worth mentioning that good results were also obtained when applying this procedure to the previous case. To further evaluate the performance of the ALS method, we generate a new simulated dataset considering the same covariance matrices as above (Q_{sim}, R_{sim}), but with the addition of two off-diagonal terms to the measurement noise covariance matrix. The following terms, associated with temperature measurements of the polymerization process, were considered in this case: $R_{\text{sim}}(15,17) = 55$; $R_{\text{sim}}(16,7) = 40$. Also, both the diagonal and the full versions of the ALS technique are implemented to estimate the covariances. For the diagonal estimation, in addition to losing the information related to the off-diagonal terms, the magnitude of the RE between the covariance estimates and their actual values was high for some cases. Specifically, the RE for the estimates of $Q(2,2)$, $Q(3,3)$, $R(7,7)$, and $R(16,16)$, all temperature-related variables, obtained were 8.13%, 100% (estimate of 0 magnitude), 36.38%, and 48.92%, respectively. However, for the full ALS technique, estimates for the off-diagonal terms of the same order of magnitude of the corresponding terms used to generate the data were obtained, $R_{\text{full}}(15,17) = 25.12$ and $R_{\text{full}}(16,7) = 22.24$. Additionally, for comparison, the RE for estimates of $Q(2,2)$, $Q(3,3)$, $R(7,7)$, and $R(16,16)$ were also calculated and all were shown to be lower than the ones for the diagonal case with magnitudes of 1.77%, 10.49%, 10.62%, and 27.67%, respectively. The differences between the REs for the two methods associated with the other process variables were not significant. Therefore, this case study shows that the full-dimensional ALS technique is able to estimate the covariances more accurately than the diagonal method for the generated dataset.

Finally, although this paper addresses the implementation of the developed method using simulated datasets, the application of this procedure using datasets from an industrial polymerization process is currently under investigation. Specifically, future studies that consider full-dimensional covariance matrices (Q, R) associated with this industrial example will be performed and the ALS estimation performance will be evaluated when the process model assumes full-dimensional versus diagonal covariance matrices. Another analysis worth pursuing for such a process is to quantify how often the covariance matrices change in the industrial dataset 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 [33].

V. Conclusion

This paper presented a method to design nonlinear state estimators for weakly observable systems. This design method included a new LTV-ALS procedure to estimate noise covariances from operating data for nonlinear systems and a novel method for model reduction in filters based on the Schmidt–Kalman filtering approach. The LTV-ALS technique was shown to give good covariance estimates from data simulated with known noise statistics. As suggested by the results in Section III, physical models for polymerization processes may be overly complex considering the available measurements; they may contain many unobservable and weakly observable modes. Overly complex disturbance models for weakly observable system models must be avoided because they lead to ill-conditioned covariance estimation and state estimation. To address this issue, the design of the reduced-order EKF to estimate only the strongly observable system states was presented. This approach also provides guidelines for industrial practitioners when designing filters for weakly observable systems, including polymerization processes. The main contributions of this paper are in the design of state estimators for weakly observable systems, nonlinear estimation using physical models, nonlinear covariance estimation from data, and building low-complexity disturbance models for nonlinear systems. Finally, the improvement in the estimator performance implies cost benefits in the implementation of advanced control schemes (see for example [33] for numerical values) when the ALS technique is used to estimate the noise covariances.

References

[1] S. A. Dadebo, M. L. Bell, P. J. McLellan, and K. B. McAuley, "Temperature control of industrial gas phase polyethylene reactors," *J. Process Control*, vol. 7, no. 2, pp. 83–95, 1997.

[2] A. Gani, P. Mhaskar, and P. D. Christofides, "Fault-tolerant control of a polyethylene reactor," *J. Process Control*, vol. 17, no. 5, pp. 439–451, 2007.

[3] K. B. McAuley, A. E. MacGregor, and A. E. Hamielec, "A kinetic model for industrial gas-phase ethylene copolymerization," *AIChE J.*, vol. 36, no. 6, pp. 837–850, 1990.

[4] F. V. Lima and J. B. Rawlings, "Nonlinear stochastic modeling to improve state estimation in process monitoring and control," *AIChE J.*, vol. 57, no. 4, pp. 996–1007, 2011.

[5] B. J. Odelson, M. R. Rajamani, and J. B. Rawlings, "A new autocovariance least-squares method for estimating noise covariances," *Automatica*, vol. 42, no. 2, pp. 303–308, Feb. 2006.

[6] M. R. Rajamani and J. B. Rawlings, "Estimation of the disturbance structure from data using semidefinite programming and optimal weighting," *Automatica*, vol. 45, no. 1, pp. 142–148, 2009.

[7] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1970.

[8] A. Alessandri, M. Baglietto, and G. Battistelli, "Moving-horizon state estimation for nonlinear discrete-time systems: New stability results and approximation schemes," *Automatica*, vol. 44, no. 7, pp. 1753–1765, Jul. 2008.

[9] C. V. Rao and J. B. Rawlings, "Constrained process monitoring: Moving-horizon approach," *AIChE J.*, vol. 48, no. 1, pp. 97–109, Jan. 2002.

[10] C. V. Rao, J. B. Rawlings, and D. Q. Mayne, "Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations," *IEEE Trans. Autom. Control*, vol. 48, no. 2, pp. 246–258, Feb. 2003.

[11] D. G. Robertson and J. H. Lee, "On the use of constraints in least squares estimation and control," *Automatica*, vol. 38, no. 7, pp. 1113–1124, 2002.

[12] D. G. Robertson, J. H. Lee, and J. B. Rawlings, "A moving horizon-based approach for least-squares state estimation," *AIChE J.*, vol. 42, no. 8, pp. 2209–2224, Aug. 1996.

[13] J. B. Rawlings and B. R. Bakshi, "Particle filtering and moving horizon estimation," *Comput. Chem. Eng.*, vol. 30, nos. 10–12, pp. 1529–1541, 2006.

[14] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*. New York: Wiley, 1997.

[15] S. F. Schmidt, "Application of state-space methods to navigation problems," in *Advances in Control Systems*, vol. 3, C. T. Leondes, Ed. New York: Academic, 1966, pp. 293–340.

[16] J. A. Farrell, *Aided Navigation: GPS with High Rate Sensors*. New York: McGraw-Hill, 2008.

[17] J. A. Farrell and M. Barth, *The Global Positioning System and Inertial Navigation*. New York: McGraw-Hill, 1998.

[18] B. A. Francis and W. M. Wonham, "The internal model principle for linear multivariable regulators," *J. Appl. Math. Optim.*, vol. 2, no. 2, pp. 170–195, 1975.

[19] B. A. Francis and W. M. Wonham, "The internal model principle of control theory," *Automatica*, vol. 12, no. 5, pp. 457–465, 1976.

[20] G. Pannocchia and J. B. Rawlings, "Disturbance models for offset-free MPC control," *AIChE J.*, vol. 49, no. 2, pp. 426–437, 2003.

[21] M. R. Rajamani, J. B. Rawlings, and S. J. Qin, "Achieving state estimation equivalence for misassigned disturbances in offset-free model predictive control," *AIChE J.*, vol. 55, no. 2, pp. 396–407, Feb. 2009.

[22] N. R. Kristensen, H. Madsen, and S. B. Jørgensen, "Parameter estimation in stochastic grey-box models," *Automatica*, vol. 40, no. 2, pp. 225–237, Feb. 2004.

[23] E. D. Sontag, *Mathematical Control Theory*, 2nd ed. New York: Springer-Verlag, 1998.

[24] A. Graham, *Kronecker Products and Matrix Calculus with Applications*. West Sussex, U.K.: Ellis Horwood, 1981.

[25] W. H. Steeb, *Kronecker Product of Matrices and Applications*. Mannheim, Germany: B.I.-Wissenschaftsver, 1991.

[26] G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications*. San Francisco, CA: Holden-Day, 1968.

[27] A. C. Aitken, "On least squares and linear combinations of observations," *Proc. Royal Soc. Edingburgh*, vol. 55, section A, pp. 42–48, 1935.

[28] M. R. Rajamani, "Data-based techniques to improve state estimation in model predictive control," Ph.D. thesis, Dept. Chem. Eng., Univ. Wisconsin-Madison, Madison, Oct. 2007.

[29] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 49–95, Mar. 1996.

[30] W. L. Brogan, *Modern Control Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1991.

[31] R. Hermann and A. J. Krener, "Nonlinear controllability and observability," *IEEE Trans. Autom. Control*, vol. 22, no. 5, pp. 728–740, Oct. 1977.

[32] C. E. de Souza, M. R. Gevers, and G. C. Goodwin, "Riccati equations in optimal filtering of nonstabilizable systems having singular state transition matrices," *IEEE Trans. Autom. Control*, vol. 31, no. 9, pp. 831–838, Sep. 1986.

[33] B. J. Odelson, A. Lutz, and J. B. Rawlings, "The autocovariance least-squares method for estimating covariances: Application to model-based control of chemical reactors," *IEEE Trans. Control Syst. Technol.*, vol. 14, no. 3, pp. 532–541, May 2006.

Fernando V. Lima received the B.S. degree from the University of São Paulo, São Paulo, Brazil, in 2003, and the Ph.D. degree from Tufts University, Medford, MA, in 2007, both in chemical engineering.

He has been a Research Associate with the University of Wisconsin, Madison, and a Post-Doctoral Associate with the University of Minnesota, Minneapolis. He will join the Faculty of chemical engineering at West Virginia University, Morgantown, as an Assistant Professor, in 2013. His current research interests include process design and operability, model-based control and optimization, state estimation and process identification, and sustainable energy systems.

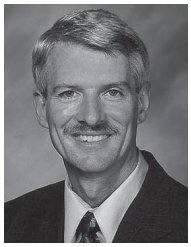
Murali Rajamani was born in Coimbatore, India. He received the Bachelor's degree in chemical engineering from the Institute of Chemical Technology, Mumbai University, Mumbai, India, in 2002, and the Ph.D. degree in chemical engineering from the University of Wisconsin, Madison, in 2007, specializing in statistics, optimization, and predictive control. He is currently pursuing the M.B.A. degree with the Booth School of Business, University of Chicago, Chicago, IL, specializing in analytical finance and econometrics.

He held various positions with BP, Naperville, IL, including in research, advanced control, operations, and technology strategy, from 2007 to 2011. His current research interests include model predictive control, state estimation, estimation of noise statistics, nonlinear control, and applications of semidefinite optimization in control and estimation.

Tyler Soderstrom received the Bachelor's degree in chemical engineering from the University of Minnesota, Minneapolis, in 1996, and the Master's and Ph.D. degrees in chemical engineering from the University of Texas at Austin, Austin, in 1998 and 2001, respectively.

He joined ExxonMobil Chemical Company, Baytown, TX, as an Advanced Control Engineer. He has held various positions with a focus on linear and nonlinear control application development and was also a Section Supervisor for a control engineering group. Currently, he is a Staff Engineer with the Core Engineering Group, ExxonMobil Chemical Company where he leads the nonlinear control program and provides worldwide support for nonlinear control applications. He maintains his connection to academia through active participation in joint academic and industrial consortia.

Dr. Soderstrom is a member of the American Institute of Chemical Engineers.



James B. Rawlings (F'12) received the B.S. degree from the University of Texas, Austin, and the Ph.D. degree from the University of Wisconsin, Madison, both in chemical engineering.

He spent a year with the University of Stuttgart, Stuttgart, Germany, as a NATO Post-Doctoral Fellow before joining the Faculty of the University of Texas. He moved to the University of Wisconsin in 1995 and is currently the Paul A. Elfers Professor and W. Harmon Ray Professor of chemical and biological engineering and the Co-Director of Texas–Wisconsin–California Control Consortium. He has published numerous research articles and co-authored two textbooks, titled *Model Predictive Control: Theory and Design* (David Mayne, 2009) and *Chemical Reactor Analysis and Design Fundamentals* (John Ekerdt, 2004). His current research interests include chemical process modeling, molecular-scale chemical reaction engineering, monitoring and control, nonlinear model predictive control, and moving horizon state estimation.

Prof. Rawlings is a fellow of the American Institute of Chemical Engineers.