Computers and Chemical Engineering 35 (2011) 71–83

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

A real-time algorithm for moving horizon state and parameter estimation

Peter Kühl[a,*], Moritz Diehl[c], Tom Kraus[b], Johannes P. Schlöder[b], Hans Georg Bock[b]

[a] BASF SE, Ludwigshafen, Germany
[b] Interdisciplinary Center for Scientific Computing, University of Heidelberg, Germany
[c] Electrical Engineering Department (ESAT), K.U. Leuven, Belgium

ARTICLE INFO

Article history:
Received 6 May 2008
Received in revised form 8 February 2010
Accepted 12 July 2010
Available online 11 August 2010

Keywords:
State estimation
Parameter estimation
Moving horizon estimation
Differential-algebraic equations
Real-time optimization
Ternary distillation
Tennessee–Eastman process

ABSTRACT

A moving horizon estimation (MHE) approach to simultaneously estimate states and parameters is revisited. Two different noise models are considered, one with measurement noise and one with additional state noise. The contribution of this article is twofold. First, we transfer the real-time iteration approach, developed in Diehl et al. (2002) for nonlinear model predictive control, to the MHE approach to render it real-time feasible. The scheme reduces the computational burden to one iteration per measurement sample and separates each iteration into a preparation and an estimation phase. This drastically reduces the time between measurements and computed estimates. Secondly, we derive a numerically efficient arrival cost update scheme based on one single QR-factorization. The MHE algorithm is demonstrated on two chemical engineering problems, a thermally coupled distillation column and the Tennessee Eastman benchmark problem, and compared against an Extended Kalman Filter. The CPU times demonstrate the real-time applicability of the suggested approach.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

In control engineering one often desires to have full knowledge of the current process state variables and, possibly, also of unknown process parameters. Retrieving these states and parameters from online measurements, possibly with the aid of an a priori model of the system, is generally referred to as state and parameter estimation. For linear systems, this task is largely solved and powerful tools such as the Kalman Filter exist (Gelb, 1974). The situation becomes more difficult for nonlinear systems. Here, most methods are extensions of linear state estimators, such as the Extended Kalman Filter described in Becerra, Roberts, and Griffiths (2001) for the class of differential-algebraic systems treated here. Further examples are observers in normal-form (Bestle & Zeitz, 1983) or high-gain-observers (Tournambe, 1992).

An excellent overview of the status quo in nonlinear state estimation is given in Daum (2005) and in Rawlings and Bakshi (2006). A comprehensive introductory overview on nonlinear state estimation is provided in the third chapter of Muske and Edgar (1997). While most classic methods rely on a recursive formulation to overcome the typical "curse of dimensionality"[1] of full-information estimators, moving horizon state estimation (MHE) is an optimization-based method that works on a horizon or "window" covering a limited number of past measurements. It has been thoroughly analyzed for unconstrained nonlinear systems (Michalska & Mayne, 1995), as well as for constrained linear (Rao, Rawlings, & Lee, 2001) and nonlinear systems (Rao, Rawlings, & Mayne, 2003; Robertson & Lee, 1995).

Both, historically and conceptually, MHE can be understood as the dual of model predictive control (MPC) and the development of both techniques has been strongly interconnected.

In MPC, a constrained optimal control problem on a finite horizon is solved for the dynamic system under consideration. Starting point is the current system state, and the solution is computed based on a model describing the future behavior of the system. A first part of the solution is applied to the process, then the optimal control problem is solved again with updated state information and a shifted horizon.

Because of the computational complexity of *nonlinear* model predictive control (NMPC), this approach has for a long time been considered an interesting theoretical concept rather than a practical control scheme. Recent improvements in computational power and solution algorithms, however, have helped NMPC to become a viable option for the control of complex processes described by nonlinear differential-algebraic equations. A good overview of the current status in NMPC research can be found in Findeisen, Allgöwer, and Biegler (2006).

NMPC and MHE share the receding horizon approach and the fact that a dynamic optimization problem repeatedly has to be

* Corresponding author.
E-mail address: peter.kuehl@basf.com (P. Kühl).

[1] We refer to the growing number of measurements in full-information estimators. The curse of dimensionality associated with the number of states to be estimated is yet another issue.

0098-1354/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2010.07.012

solved online. Advantages of the MHE formulation are the explicit consideration of state and parameter constraints, the optimality of the estimates in a least-squares sense, and the proven stability properties as shown in Rao et al. (2001, 2003).

We see another advantage of MHE in the fact that disturbances in the form of unknown and slowly time-varying parameters can be estimated along with the states in a consistent way by adding them as single degrees of freedom to the optimization problem. This is in contrast to many other estimation approaches where parameters have to be reformulated as additional states.

The practical necessity of estimating process parameters together with state variables is less intuitive and deserves a remark: NMPC (as well as other model-based control methods) heavily relies on an accurate process model. Model-plant mismatch will deteriorate the optimal control solution and may lead to steady-state offsets and to oscillations or even instability of the closed-loop. This problem can partially be addressed by using a suitable disturbance model and adapting model parameters to current measurements. For more details see Muske and Badgwell (2002), Pannocchia and Rawlings (2003) and Faanes and Skogestad (2005).

The need for real-time optimal control has stimulated intensive research and a large arsenal of fast algorithms for NMPC is now available (Cannon, 2004; Diehl et al., 2002; Martinsen, Biegler, & Foss, 2004; Schäfer, Kuehl, Diehl, Schlöder, & Bock, 2007; Zavala, Laird, & Biegler, 2006). Additional progress is visible as NMPC enters the world of mechanical systems with sampling times in the range of milliseconds as demonstrated, e.g., in Ferreau, Lorini, and Diehl (2006).Fewer results have been published, however, in the field of optimization-based *online state and parameter estimation*, and only a few publications treat fast numerical (real-time) algorithms. Early work in this direction was done by Ohtsuka and Fujii (1996) and Tenny and Rawlings (2002). Interesting algorithms are presented in Valdes-Gonzalez, Flaus, and Acuna (2003) for a globally convergent MHE (not yet real-time feasible, however) and in Jorgensen, Rawlings, and Jorgensen (2004) for linear systems. In Kang (2006), an algorithm based on backwards single-shooting is presented along with a stability analysis that even considers inexact solution of the model equations. Only very recently, a convincing real-time algorithm based on collocation has been presented in Zavala, Laird, and Biegler (2007). Despite many efforts in the field, Rawlings and Bakshi (2006) state that "*computational complexity remains a significant research challenge for MHE researchers*".

This challenge is tackled in the present work leading to an algorithm which allows for real-time applicability of MHE even for large-scale chemical engineering processes. The work is inspired by Diehl et al. (2002, 2005). A similar version of the article has been published by the same authors in German (Diehl, Kuehl, Bock, & Schlöder, 2006).

The article proceeds as follows: the general MHE formulation is given in Section 2, followed by the derivation of an update scheme for the arrival costs. The main feature of this article, the MHE real-time iteration for a fast numerical solution of the MHE problem is explained in Section 3. This section includes a brief introduction to multiple shooting and a summary of the real-time algorithm. The proposed method is demonstrated in Section 4 by applying it to a coupled distillation column (Section 4.1) and to the Tennessee–Eastman benchmark problem (Section 4.2). The article ends with a short summary and outlook.

1.1. Problem formulation

We consider dynamic systems modeled by a set of differential-algebraic equations (DAEs):

$$\dot{x}(t) = f(x(t), z(t), u(t), p), \qquad x(t_0) = x_0, \tag{1a}$$

$$0 = g(x(t), z(t), u(t), p), \tag{1b}$$

$$y(t) = h(x(t), z(t), p), \tag{1c}$$

with differential states $x \in \mathbb{R}^{n_x}$, algebraic states $z \in \mathbb{R}^{n_z}$, controls $u \in \mathbb{R}^{n_u}$, parameters $p \in \mathbb{R}^{n_p}$, and outputs $y \in \mathbb{R}^{n_h}$. It is assumed that the matrix $\partial g/\partial z$ is regular, i.e. that the DAE system is of index 1. We note that a DAE system of higher index can often be transformed to this form by index reduction techniques.

The continuous time model is then transformed to discrete-time with the help of the analytic solution or numerical integration. Our MHE scheme builds on the resulting discrete-time model:

$$x_{k+1} = F(x_k, z_k, u_k, p) + w_k, \qquad x_0 = x(t_0), \tag{2a}$$

$$0 = g(x_k, z_k, u_k, p), \tag{2b}$$

$$y_k = h(x_k, z_k, p) + v_k. \tag{2c}$$

Here, index k denotes the samples taken at times t_k. For ease of notation only, the controls $u(t)$ are assumed piecewise constant over the time intervals $t \in [t_k, t_{k+1}]$, denoted by u_k. Measurement errors in the measured outputs at samples k are expressed by $v_k \in \mathbb{R}^{n_h}$. To account for unknown disturbances on the system states, a simple additive state noise term with $w_k \in \mathbb{R}^{n_x}$ can be added. Based on the above process model, the MHE problem is formulated: At time t_k we consider a horizon containing M measurements $\{y_{k-M+1}, \ldots, y_k\}$ taken at times $t_{k-M+1} < \cdots < t_k$. The length of the estimation horizon is $T_E = t_k - t_{k-M+1}$ and we set $k - M + 1 =: L$. Depending on the presence of additive state noise in the model, we distinguish between two different MHE formulations.

1.2. Output noise MHE

For this formulation the dynamic model is *assumed* to be perfectly known with no unmeasured disturbances perturbing the states. Therefore, the state noise term w_k in Eq. (2a) is omitted. The MHE problem is now formulated as the following constrained least-squares optimization problem:

$$\min_{\substack{x_j, z_j, \\ p}} \left(\left\| \begin{matrix} x_L - \bar{x}_L \\ p - \bar{p}_L \end{matrix} \right\|^2_{P_L} + \sum_{j=L}^{k} \|y_j - h(x_j, z_j, p)\|^2_{V_j} \right), \tag{3a}$$

$$\text{s.t.} \quad \left.\begin{aligned} x_{j+1} &= F(x_j, z_j, u_j, p), \\ 0 &= g(x_j, z_j, u_j, p), \\ x_{j,\min} &\le x_j \le x_{j,\max}, \\ z_{j,\min} &\le z_j \le z_{j,\max}, \end{aligned}\right\} \quad j = L, \ldots, k \\ p_{\min} \le p \le p_{\max}. \tag{3b}$$

The first term of the cost function in Eq. (3) (and also in Eq. (4)) is typically called "arrival cost" and is well-known to play an important role for MHE. It will be discussed in more detail below when our approach to determine $(\bar{x}_L, \bar{p}_L)$ and P_L is introduced.

2. Output and state noise MHE

In this formulation, both measurement noise and state noise are considered. In essence, this adds additional degrees of freedom to the optimization problem and the MHE problem becomes finding states, parameters and state noise terms that solve the following constrained least-squares optimization problem:

$$\min_{\substack{x_j, z_j, \\ p, w_j}} \left(\left\| \begin{matrix} x_L - \bar{x}_L \\ p - \bar{p}_L \end{matrix} \right\|^2_{P_L} + \sum_{j=L}^{k} \|y_j - h(x_j, z_j, p)\|^2_{V_j} + \sum_{j=L}^{k-1} \left\| w_j \right\|^2_W \right), \tag{4a}$$

$$\text{s.t.}\quad \begin{aligned} & x_{j+1} = F(x_j, z_j, u_j, p) + w_j, \quad j = L, \ldots, k-1, \\ & \left.\begin{aligned} & 0 = g(x_j, z_j, u_j, p), \\ & x_{j,\min} \le x_j \le x_{j,\max}, \\ & z_{j,\min} \le z_j \le z_{j,\max}, \\ & w_{j,\min} \le w_j \le w_{j,\max}, \\ & p_{\min} \le p \le p_{\max}. \end{aligned} \right\} \quad j = L, \ldots, k \end{aligned} \tag{4b}$$

Slightly different from conventional notation, we write $\|a\|_A^2 = a^T A^T A a$. The weighting matrices in both cost functions can be interpreted as inverses of covariance matrices. So we have $P_L = Q_0^{-(1/2)}$ where Q_0 is a block-diagonal matrix with the two symmetric blocks Q_0^x and Q_0^p; $V_j = R_j^{-(1/2)}$; $W = Q^{-(1/2)}$.

Solving either MHE problem at time k yields a state estimate and an estimate of the free parameters. Let us now suppose that the following assumptions hold:

- The initial value $x_0 = x(0)$ is a normally distributed random variable with covariance matrix Q_0^x.
- The parameter vector p is a normally distributed random variable with covariance matrix Q_0^p.
- The noise sequences $\{w_k\}$ and $\{v_k\}$ are independent, normally distributed random variables with covariances Q and R_k.

Under these assumptions it is well-known that estimates resulting from the solution of the unconstrained MHE problem can stochastically be interpreted as maximum-likelihood estimates. For constrained noise terms, truncated normal distributions provide meaningful statistics, while constraints on the states may introduce an acausality which does not allow for a meaningful stochastic interpretation of the estimation results anymore, see Robertson, Lee, and Rawlings (1996). The case of non-Gaussian distributions is reviewed in and. Note that in principle, MHE can deal with any given distribution provided that the optimization problem is formulated as minimizing the negative logarithm of the noise density functions and the arrival cost is chosen as the negative logarithm of the conditional density for x_L given all measurements up to time $L-1$. In practice, however, these conditional densities are not at hand. Haseltine and Rawlings (2005) have demonstrated that MHE designed for normal distributions is likely to fail for multi-modal density functions. Very recently, Ungarala (2009) has suggested to use nonparametric arrival cost terms to cope with non-Gaussian distributions. This approach, however, suffers from a relatively high computational demand.

The formulations proposed in this article may still deliver meaningful results for non-Gaussian distributions or $\{w_k\}$ and $\{v_k\}$ being dependent variables as long as the weighting matrices are positive definite. However, nothing can then be said a priori about the statistical nature of the estimates.

Apart from the noise characteristics, other aspects may play a role in practical applications, e.g., non-equidistant sampling times, continuous measurements, delayed measurements, and (nonlinear) path constraints on the system states. These aspects have all been omitted for brevity but can, in principle, be incorporated into the proposed algorithm.

In the next section, an efficient alternative to update the arrival cost required by both MHE formulations is derived.

2.1. Efficient update of the arrival cost

The receding horizon window approach limits the size of the optimization problem to overcome the "curse of dimensionality" of a full-information estimator. To summarize information contained in measurements before $j = L$, an arrival cost term has been proposed and further developed by a number of authors (Muske, Rawlings, & Lee, 1993; Rao et al., 2001; Robertson et al., 1996; Tenny, 2002). It was shown that a good choice of the arrival cost allows to approximate the full-information estimation and is key to stability of MHE. Typically, the arrival cost is derived from dynamic programming arguments and uses Kalman Filter-based updates for the weighting matrix P_L. Here, the arrival cost is motivated in a slightly different way, automatically leading to a convenient update formulation.

Ideally, one would like to use the exact arrival cost

$$\min \left(\sum_{j=-\infty}^{L} \|y_j - h(x_j, z_j, p)\|_{V_j}^2 + \sum_{j=-\infty}^{L-1} ||w_j||_W^2 \right) \quad \text{s.t. (2).}$$

Note that we have omitted any constraints for the time being. With the exception of linear systems, these exact costs cannot be expressed analytically. Therefore, the goal now is to approximate the arrival cost by a quadratic term that is updated prior to each new horizon shift. Given the solution $x^*(\cdot)$, $z^*(\cdot)$, p^* of the MHE problem on interval $[t_L, t_k]$ as well as data from the previous MHE problem, the new arrival cost data $(\bar{x}_{L+1}, \bar{p}_{L+1})$ and $P_{L+1}a$ for the subsequent MHE problem on interval $[t_{L+1}, t_{k+1}]$ is calculated.

By hypothetically considering the MHE problem on the increased interval $[t_L, t_{k+1}]$ one observes that the information contained in $[t_L, t_{L+1}]$ – consisting of the *old* arrival cost, the first measurement y_L, and the state noise w_L – must be summarized in a *new* arrival cost term for the interval $[t_{L+1}, t_{k+1}]$. It is indeed possible to formulate an MHE problem on $[t_{L+1}, t_{k+1}]$ fully equivalent to the hypothetical problem by introducing a specific nonlinear function $C(x_{L+1}, p_{L+1})$ as the new arrival cost. Here and in the following, x_{L+1} and p_{L+1} denote the values $x(t_{L+1})$ and p of the new problem on $[t_{L+1}, t_{k+1}]$. Standard dynamic programming arguments establish that the ideal arrival cost C must be

$$\begin{aligned} C(x_{L+1}, p_{L+1}) = \min_{x_L, p_L} & \left(\left\| \begin{matrix} x_L - \bar{x}_L \\ p_L - \bar{p}_L \end{matrix} \right\|_{P_L}^2 + \|y_L - h(x_L, p_L)\|_{V_L}^2 + \left\| \begin{matrix} w_L \\ w_L^p \end{matrix} \right\|_{\bar{W}_L}^2 \right), \\ \text{s.t.}\quad & w_L = x_{L+1} - F(x_L, u_L, p_L), \\ & w_L^p = p_{L+1} - p_L. \end{aligned} \tag{5}$$

The algebraic state vector z_L does not play a role for deriving the arrival cost update approach and has been omitted for the sake of a simpler presentation. This is conceptually justified because z_L can always be expressed as a function of x_L, u_L, and p by solving Eq. (2b).

Note that "parameter noise" w_L^p has been introduced in addition to the process noise w_L. This allows to estimate time-varying parameters despite the model assumption of constant parameters as in Eq. (3). The weighting matrix $\bar{W}_L = \text{diag}(Q, Q^p)^{-(1/2)}$ is created as the inverse of a block-diagonal matrix with block Q (introduced before as the state noise covariance matrix) and the parameter noise covariance Q^p.

We now aim at approximating this "ideal" arrival cost C by a function of the form:

$$C'(x_{L+1}, p_{L+1}) = \text{const} + \left\| \begin{matrix} x(t_{L+1}) - \bar{x}_{L+1} \\ p - \bar{p}_{L+1} \end{matrix} \right\|_{P_{L+1}}^2. \tag{6}$$

Since the constant term does not play a role in the MHE optimization problem, we would then have achieved our aim to summarize all information in $[t_L, t_{L+1}]$ in a quadratic term and could set up a new MHE problem for the shifted horizon.

In what follows, the quadratic approximation of the ideal arrival cost C is derived. We denote the solution of the DAE system describing the system on the interval $t \in [t_L, t_{L+1}]$ with initial value x_L and parameter p_L by $x(t_{L+1}; x_L, p_L)$. The corresponding algebraic states are $z(t_{L+1}; x_{L+1})$. Observing that C cannot be expressed analytically because of the two nonlinear functions $x(t_{L+1}; x_L, p_L)$ and $h(x_L, z_L,$

p_L), we linearize them around the best available estimate ($x^*(t_L)$, p^*).[2] This yields

$$x(t_{L+1}; x_L, p_L) \approx \tilde{x} + \underbrace{\frac{dx(t_{L+1}; x^*(t_L), p^*)}{dx(t_L)}}_{=:X_x} x_L + \underbrace{\frac{dx(t_{L+1}; x^*(t_L), p^*)}{dp}}_{=:X_p} p_L,$$

with

$$\tilde{x} := x^*(t_{L+1}; x^*(t_L), p^*) - X_x x^*(t_L) - X_p p^*,$$

and, analogously,

$$h(x_L, z_L, p_L) \approx \tilde{h} + H_x x_L + H_p p_L.$$

The linearization transforms problem (5) into an analytically solvable least-squares problem of the form

$$\min_{x_L, p_L} \left\| \begin{pmatrix} P_L \begin{pmatrix} x_L - \bar{x}_L \\ p_L - \bar{p}_L \end{pmatrix} \\ V_L(y_L - \tilde{h} - H_x x_L - H_p p_L) \\ \bar{W}_L \begin{pmatrix} x_{L+1} - \tilde{x} - X_x x_L - X_p p_L \\ p_{L+1} - p_L \end{pmatrix} \end{pmatrix} \right\|_2^2 .$$

The problem with free variables $x_L, p_L, x_{L+1}, p_{L+1}$ can be transformed via QR-factorization of the matrix:

$$\left(\begin{array}{c} P_L \\ -(V_L H_x | V_L H_p) \\ -\bar{W}_L \begin{pmatrix} X_x & X_p \\ 0 & \mathbb{I} \end{pmatrix} \end{array} \middle| \begin{array}{c} 0 \\ 0 \\ \bar{W}_L \end{array} \right) = \mathcal{Q} \begin{pmatrix} \mathcal{R}_1 & \mathcal{R}_{12} \\ 0 & \mathcal{R}_2 \\ 0 & 0 \end{pmatrix}$$

into an equivalent problem

$$\min_{x_L, p_L} \left\| \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix} + \begin{pmatrix} \mathcal{R}_1 & \mathcal{R}_{12} \\ 0 & \mathcal{R}_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_L \\ p_L \\ x_{L+1} \\ p_{L+1} \end{pmatrix} \right\|_2^2 .$$

It has the analytic solution

$$C'(x_{L+1}, p_{L+1}) = \|\rho_3\|_2^2 + \left\| \rho_2 + \mathcal{R}_2 \begin{pmatrix} x_{L+1} \\ p_{L+1} \end{pmatrix} \right\|_2^2,$$

and for the arrival cost update we obtain

$$P_{L+1} := \mathcal{R}_2, \qquad \begin{pmatrix} \bar{x}_{L+1} \\ \bar{p}_{L+1} \end{pmatrix} := -\mathcal{R}_2^{-1} \rho_2. \tag{7}$$

Remark 1. For all vectors $v \in \mathbb{R}^{n_x + n_p}$ we have that $\|v\|_{P_{L+1}}^2 \leq \|v\|_{\bar{W}_L}^2$ (which immediately follows from the identity $\mathcal{R}_{12}^T \mathcal{R}_{12} + \mathcal{R}_2^T \mathcal{R}_2 = \bar{W}_L^T \bar{W}_L$). This means that the a priori initial value $\bar{x}_{L+1}$ and the parameter $\bar{p}_{L+1}$ can only be determined within the limits given by the noise with covariance $Q = \bar{W}_L^{-2}$, which guarantees that the influence of past information cannot grow excessively.

Remark 2. If only the most current measurement is used, i.e. $M = 1$, the described MHE formulation is equivalent to the Extended Kalman Filter (EKF). This is obtained almost immediately by comparing the QR-factorization to the square-root formulation of the Kalman Filter. As a consequence, the arrival cost update in our MHE algorithm inherits all beneficial numerical properties known from the square-root Kalman Filter formulation.

[2] This corresponds to the smoothing update mentioned in Rao et al. (2001).

3. Numerical solution of the MHE problem

In this section we present the *real-time iteration* (RTI) approach to efficiently solve MHE problems. This algorithmic concept essentially reduces the computational burden to only one iteration of a sequential quadratic programming (SQP) method per sample. Rather than fully solving each single MHE problem to convergence while losing precious time working on ageing measurement data, the solution of the underlying least-squares problem is coupled with the evolving process. The particular structure of RTI also allows one to split each iteration into two phases: a preparation phase carried out during two sampling instants and an estimation phase that delivers state estimates almost instantaneously after the latest measurement enters the MHE problem. The RTI concept was originally proposed in (Bock, Diehl, Leineweber, & Schlöder, 2000) and has been integrated into an NMPC algorithm by Diehl et al. (2002). It builds on the framework of multiple shooting for the direct solution of optimal control problems (Bock & Plitt, 1984). Before stating the *MHE real-time iteration algorithm*, we briefly review essential ingredients of our numerical solution approach. For a more comprehensive introduction to direct multiple shooting we refer the reader to the two seminal articles by Leineweber, Bauer, Schäfer, Bock, and Schlöder (2003).

3.1. A brief introduction to direct multiple shooting

Several methods exist to solve dynamic optimization problems with underlying DAE systems, see, e.g., Binder et al. (2001) for an overview. Direct multiple shooting used in this work has a long tradition in off-line parameter estimation (Bock, 1983). It belongs to the group of direct methods and is characterized by the simultaneous solution of optimization problem and underlying DAE system. Direct multiple shooting aims at solving dynamic optimization problems of the form

$$\begin{aligned} \min_{u \in \mathcal{U}} \; & F(x(t), z(t), u(t), p) \\ \text{s.t.} \; & \dot{x} = f(x(t), z(t), u(t), p), \qquad x(t_0) = x_0, \quad t \in [t_0, t_0 + T], \\ & 0 = g(x(t), z(t), u(t), p) \\ & 0 \geq r(x(t), z(t), u(t), p). \end{aligned} \tag{8}$$

First, a multiple shooting time grid with $m+1$ grid points is introduced on $[t_0, t_0 + T]$: $t_0 = \tau_0 < \tau_1 < \cdots < \tau_m = t_0 + T$. The control space is then parameterized with a finite number of control parameters q on the multiple shooting grid to transfer the infinite-dimensional optimal control problem (8) to a finite one in the control variables. The most popular choice is a piecewise constant representation of the control function $u(t)$ with control parameters q_j in each shooting interval:

$$u(\tau) = q_j, \quad \tau \in I_j := [\tau_j, \tau_{j+1}], \quad j = 0, \ldots, m-1.$$

Next, the state variables are discretized along the multiple shooting grid. Start values s_j^x, s_j^z for the differential and algebraic states are used to define initial value problems on each multiple shooting interval I_j:

$$\dot{x}_j(\tau) = f(x_j(\tau), z_j(\tau), q_j, p), \qquad 0 = g(x_j(\tau), z_j(\tau), q_j, p)$$

with initial values $x(\tau_j) = s_j^x$ and $z(\tau_j) = s_j^z$. These initial values can be chosen arbitrarily.

To obtain a solution of the original problem despite the relaxation, so-called *continuity conditions* $x(\tau_{j+1}; s_j^x, s_j^z, q_j) - s_{j+1}^x = 0$, $j = 0, 1, \ldots, m-1$ and *consistency conditions* $g(s_j^x, s_j^z, \varphi_j(\tau_j, q_j)) = 0$, $j = 0, 1, \ldots, m$ are imposed at every grid point. Here, $x(\tau_{j+1}; s_j^x, s_j^z, q_j)$ denotes the initial value problem solution at time τ_{j+1} with initial values s_j^x, s_j^z.

After control parameterization and state discretization, one yields a structured NLP (nonlinear programming) problem formulation of (8) with new free variables $\{q_j, s_j^x, s_j^z\}$ subject to continuity and consistency conditions, initial conditions, and discretized path constraints $0 \geq \quad r(s_j^x, s_j^z, q_j, p), \quad j = 0, 1, \ldots, m$. This resulting NLP can finally be solved with, e.g., sequential quadratic programming (SQP) or interior point methods.

3.2. Direct multiple shooting solution to the MHE problem

The structure of MHE problems (3) and (4) show strong similarities with the structure obtained by the multiple shooting parameterization and discretization of a general dynamic optimization problem (8):

- The sampling instants within the estimation horizon naturally form a multiple shooting time grid.
- While $u(t)$ in the MHE problem formulation is given from the past and can be treated as constant in the optimization problem, the free variables are the states $\{x_j, z_j\}$, $j = L, \ldots, k$ to be estimated. They correspond to the direct multiple shooting variables $\{s_j^x, s_j^z\}$. In the case of output and state noise we additionally have free variables $\{w_j\}, \quad j = L, \ldots, k-1$ that correspond to the "controls" q_j in the multiple shooting formulation.
- The general cost function $F(\cdot)$ in (8) is replaced by the least-squares cost function (3a) or (4a).
- Parameters p to be estimated are added to the list as free variables that remain constant for all sampling intervals/shooting intervals within the estimation horizon.
- The discrete-time process model $x_{j+1} = F(x_j, z_j, u_j, p)$ for a sampling interval $(j, j+1)$ can be seen as an initial value problem on the corresponding shooting interval. And indeed the time discrete model is usually obtained by solving an initial value problem based on the continuous DAE system (1).

Therefore, direct multiple shooting forms a natural framework to numerically solve MHE problems (3) and (4). In detail, it transforms MHE problem (4) to a least-squares NLP of the following form:

$$\begin{aligned}
\min_{\substack{w_L, \ldots, w_k,\\ x_L, \ldots, x_k,\\ z_L, \ldots, z_k,\\ p}} \quad & \left(\left\| \begin{matrix} x_L - \bar{x}_L \\ p - \bar{p}_L \end{matrix} \right\|^2_{P_L} + \sum_{j=L}^{k} \|y_j - h(x_j, z_j, p)\|^2_{V_j} + \sum_{j=L}^{k-1} \|w_j\|^2_W \right) \\
\text{s.t.} \quad & x_{j+1} = F(x_j, z_j, u_j, p) + w_j, \quad j = L, \ldots, k-1, \\
& 0 = g(x_j, z_j, u_j, p), \quad j = L, \ldots, k, \\
& x_{j,\min} \leq x_j \leq x_{j,\max}, \quad j = L, \ldots, k, \\
& z_{j,\min} \leq z_j \leq z_{j,\max}, \quad j = L, \ldots, k, \\
& w_{j,\min} \leq w_j \leq w_{j,\max}, \quad j = L, \ldots, k, \\
& p_{\min} \leq p \leq p_{\max}.
\end{aligned}$$

Note that everything works analogously for problem (3). One then has to use the correct cost function and omit the free variables w_j. Without loss of generality we will focus on problem (4) in the following.

3.3. Generalized Gauss–Newton method

Collecting all free variables of the above least-squares problem in a single vector:

$$r_k = (x_L, z_L, x_{L+1}, z_{L+1}, \ldots, x_k, z_k, w_L, \ldots, w_{k-1}, p),$$

the transformed problem (9) can be stated short as

$$\begin{aligned}
& \min_{r_k} \|J(r_k; D_k)\|_2^2, \\
& \text{s.t.} \quad G(r_k; D_k) = 0, \qquad H(r_k; D_k) \geq 0.
\end{aligned} \tag{10}$$

Here, D_k represents the set of *input variables* to the MHE problem, i.e. all variables that have to be provided to solve for the current state and parameter estimates at sampling instant k. It comprises of the set of measurements and past controls within the estimation window and the current arrival cost parameters $\bar{x}_L$, $\bar{p}_L$, and P_L. We use the notation D_k to emphasize that each MHE problem at time t_k depends on a new set of data available at t_k.

Problem (10) is solved using the generalized Gauss–Newton (GGN) method (for details see Bock, 1981), an iterative method that makes use of linearizations of all problem functions around the current iteration value r_k^i (here, k still denotes the "process time", while $i = 0, \ldots, n$ denotes the ith iteration for the MHE problem at time k). With these linearizations a constrained quadratic program (QP) is formulated to find an increment Δr_k^i such that $r_k^{i+1} = r_k^i + \Delta r_k^i$ is the next iterate. The QP subproblem is given by

$$\min_{\Delta r_k^i} \left\| J(r_k^i; D_k) + \nabla_r J(r_k^i; D_k)^T \, \Delta r_k^i \right\|_2^2 \tag{11}$$

$$\text{s.t.} \quad G(r_k^i; D_k) + \nabla_r G(r_k^i; D_k)^T \, \Delta r_k^i = 0, \tag{12}$$

$$H(r_k^i; D_k) + \nabla_r H(r_k^i; D_k)^T \, \Delta r_k^i \geq 0. \tag{13}$$

An appealing feature of the GGN method is that the Hessian matrix needed to solve QP (11) can cheaply be computed by means of Jacobians only. Note that the GGN method, like all Newton-type methods, converges to a local optimum only.

The entire direct multiple shooting algorithm and the GGN are implemented in the software package MUSCOD-II. All necessary integration and differentiation steps are calculated with the DAE-solver DAESOL (Bauer, Bock, & Schlöder, 1999), which uses a backward differentiation formula (BDF) method. This solver also generates all required sensitivities of the state trajectories with respect to initial values and free variables.

In a classic MHE approach, problems (3) and (4) would now be solved for each time k until a pre-specified convergence criterion is met. This necessarily means repeatedly solving the sub-QP 11 while the process actually evolves. This is avoided by the real-time iteration scheme as described in the following section.

3.4. Real-time iterations for MHE

(9) The MHE real-time iteration is a direct extension of ideas developed in the context of real-time optimal control by Diehl et al. (2002). It reduces the computational burden of solving the least-squares problem at each time instant to one single generalized Gauss-Newton iteration (therefore we drop the iteration index i used in the previous section). This iteration can furthermore be separated into a preparation and an estimation phase. Similar approaches have been proposed in the literature, e.g., M'hamdi, Helbig, Abel, and Marquardt (1996), but suffer from the weaker contraction properties of the single-shooting method they are based on.

The iteration starts with the solution r_{k-1} of the previous sampling instant t_{k-1}. To initialize the approaching MHE problem at time t_k, an initial problem vector r_k^- is created by discarding the oldest information (which leaves the estimation horizon), shifting r_{k-1} and predicting the states at time t_k as a solution $x(t_k; t_{k-1}, x_{k-1}, z_{k-1}, u, p)$ of the corresponding initial value problem. Then, all but one component of the linear optimization problem (10) are pre-computed along with all necessary sensitivities (the latter is

Table 1
Algorithm: MHE real-time iteration.

Input: multiple shooting data vector

$$r_{k-1} = (\ x_{L-1}, z_{L-1}, x_L, z_L, ..., x_{k-1}, z_{k-1}, w_{L-1}, ..., w_{k-2}, p)$$

online data set

$$D_{k-1} = (\ \bar{x}_{L-1}, \bar{p}_{L-1}, P_{L-1}, y_{L-1}, V_{L-1}, u_{L-1} ..., y_{k-1}, V_{k-1}, u_{k-1}, y_k)$$

Output: current estimates $\hat{x}_k$, $\hat{z}_k$, $\hat{p}_k$,

updated r_k and updated $\bar{x}_L$, $\bar{p}_L$ for D_k

0. At initial sample k=0 provide initial guess w_0, D_0

FOR samples $k = 1, 2, \ldots$ **DO**

Preparation phase for horizon $[t_L, t_k]$ **before** t_k **:**

1. Update arrival cost data $(\bar{x}_L, \bar{p}_L)$ and P_L with eq. (7)

2a. Shift data vectors:

$$r_k^- = (x_L, z_L, x_{L+1}, z_{L+1}, ..., x_k^-, z_k^-, w_L, ..., w_{k-1}^-, p)$$

$$D_k = (\ \bar{x}_L, \bar{p}_L, P_L, y_L, V_L, ..., *, V_k; u(t), t \in [t_L, t_k])$$

where $*$ is later replaced by measurement y_k

2b. Solve model DAE on interval $[t_{k-1}, t_k]$ and set

$$x_k^- := x(t_k; x_{k-1}, z_{k-1}),\ w_{k-1}^- = 0,$$

$$z_k^- := z(t_k; x_{k-1}, z_{k-1})$$

3a. Compute all but one vector component of QP (11)

Only last component of $C(r_k^-; D_k)$, i.e.

$V_k(y_k - h(x_k, z_k, p))$ undefined because of missing y_k

3b. Compute *all* matrix components for solving QP (11)

(possible because y_k enters linearly only!)

Estimation phase for horizon $[t_{L+1}, t_k]$ **at** t_k **:**

4. Once measurement y_k becomes available

compute missing component in QP (11) and solve for Δr_k

6. Calculate $r_k = r_k^- + \Delta r_k$ and set $\hat{x}_k = x_k$, $\hat{z}_k = z_k$, $\hat{p}_k = p$

END

possible because the unknown measurement y_k enters the problem linearly). This concludes the preparation phase.

Upon arrival of a new measurement y_k, the missing component in (10) is evaluated and the problem is solved for the new solution vector r_k whose last entries z_k, z_k, and p are the desired current estimates $\hat{x}_k$, $\hat{z}_k$, and $\hat{p}_k$.

The entire algorithm is shown in Table 1.

Remark 3. Special precaution is needed in the starting phase of the algorithm, when the estimation window is smaller than the number of available measurements (i.e. when $k < M$). This detail is omitted in the algorithm as one can simply fill the window with the initial state or use a growing horizon approach.

Remark 4. The described algorithm is particularly appealing for real-time environments since it consists of a well-defined and finite number of matrix-vector operations provided that any further adaptive element, e.g., in the integrator, is limited to a maximum number of iterations. Then it is possible to derive a guaranteed

upper bound of the computation time which fulfills the formal definition of the real-time paradigm.

Remark 5. From a technical point of view, our approach falls under the generic moving horizon numerical observer framework presented in Kang (2006). There, stability of MHE is studied generally for any integrator and optimizer realization. The implementation suggested therein, however, is based on Euler's integration and an inexact Newton method with BFGS updates for the Hessian.

3.5. Problems and pitfalls of MHE real-time iterations

The real-time iteration algorithm shows remarkable robustness in simulation studies such as the two presented in Section 4. Nevertheless, a few critical remarks are due at this stage. It is known that observability of a system (as, e.g., defined in Rao et al., 2003) translates to the existence of a unique global solution to the MHE problems (3). Direct multiple shooting, however, as well as many other numerical approaches for dynamic optimization only finds *local* solutions in the vicinity of the initial guess. Without extensive analysis, nothing can be said a priori about the local or global nature of the solution and the respective convergence region. Even more difficulties with local solutions and small regions of convergence may be expected for MHE problem (4). These concerns hold even more so for parameter estimation in the MHE context.

The real-time approach delivers sub-optimal solutions to the MHE problem in the early stages of the estimation or when changes in parameters or controls occur. While sub-optimality has been addressed in the literature (e.g., Kang, 2006; Rao et al., 2003), a comprehensive analysis of the stability of the RTI-scheme is not yet available. The length of the estimation window plays a crucial role in MHE algorithms in general. So far, no systematic way to determine the optimal number of measurements M is available. The same is true for the number of shooting nodes, accuracies for the DAE solvers and many more solution parameters. Having stated these limitations and warnings, we now present two chemical engineering examples where the RTI approach worked well and delivered good performance.

4. Two case studies and comparison with EKF

The MHE algorithm is applied to two chemical process engineering problems. The simulation results are compared to those obtained by an Extended Kalman Filter (EKF). The first example is a thermally coupled distillation column for ternary separation. In the second example, the MHE algorithm is applied to the Tennessee Eastman (TE) benchmark process under conditions that include measurement *and* state noise. In both cases the EKF is implemented as suggested in Becerra et al. (2001), and reasonably tuned. The same covariance matrices are then used for the MHE as weighting matrices. The numerical methods to obtain sensitivities, predict states and compute consistent algebraic estimates are exactly the same for both MHE and EKF.

To allow for parameter estimation with the EKF, the respective parameters are formulated as additional differential states with Gaussian noise as right-hand-side and nominal parameter values as initial values (often called "integrated disturbance model").

4.1. Application to thermally coupled distillation columns

MHE and EKF are applied to estimate the states of two coupled distillation columns described by 106 differential and 371 algebraic variables. Additionally, three model parameters describing the feed flow are assumed to be unknown and have to be estimated. The column configuration is sketched in Fig. 1. The column separates

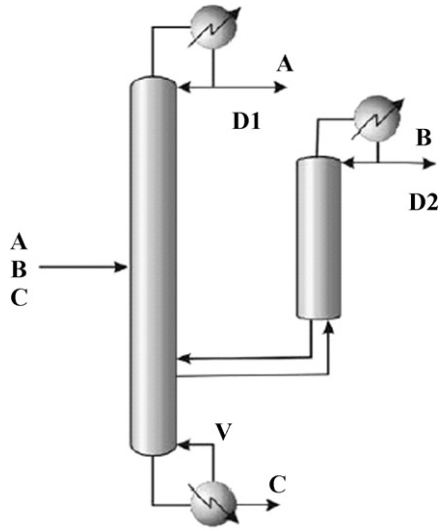


Fig. 1. Coupled distillation columns with side outlet for the ternary separation of components A, B, and C. Manipulated variables (MV) are the boil-up V and the distillates D_1 and D_2 that leave the system.

a mixture of methanol, ethanol and 1-propanol. Methanol (A) and ethanol (B) are removed overhead with the distillates of main and side column, while 1-propanol (C) is accumulated in the bottoms of the main column.

4.1.1. Model of the coupled columns

The main column consists of 42 trays (including bottoms and condenser). The side outlet is located at tray 11, the inlet at tray 21. The side column consists of 11 trays including the second condenser.

The nonlinear DAE system describes the molar concentrations of two components on all 53 trays as differential variables, while the third component's tray concentrations, all temperatures, and the internal molar flows are described by 371 algebraic variables. For modeling details we refer to (Itigin et al., 2003) and references therein. The same model also served as an example for full-state real-time NMPC in Schäfer et al. (2007).

The column is controlled in a D–V configuration with the two distillate flow rates D_1 and D_2 and the boil-up rate V serving as controls, i.e. $u = (D_1, D_2, V)^T$.

4.1.2. MHE simulation study

For the numerical study presented here we employ MHE formulation (2). It is assumed that ten trays are equipped with temperature sensors[3] and that historical data for the controls D_1, D_2 and V is available. The sampling rate is set to 200 s. The estimation horizon is set to $T_E = 800$ s, covering $m = 5$ sets of the 10 temperature measurements. We assume the temperature measurements to be corrupted by Gaussian noise with zero mean and 0.5 K standard deviation.

Process parameters to be estimated are the feed rate F and the feed concentrations $x_{A,F}$ and $x_{B,F}$. This scenario reflects variations in upstream production processes.

The lower and upper bounds for all differential states are given as $x_{\min} = 0$ and $x_{\max} = 1$ for the MHE problem. Bounds on the parameters $p = (F, x_{A,F}, x_{B,F})$ are $p_{\min} = (4 \times 10^{-2}\,\text{mol/s}, 0, 0)$ and $p_{\max} = (6 \times 10^{-2}\,\text{mol/s}, 0.5, 0.3)$. Note that the EKF algorithm is not able to take such a priori bounds into account.

[3] The positions have been chosen by plotting the steady-state temperature profiles and then determining the separation zones by visual inspection.

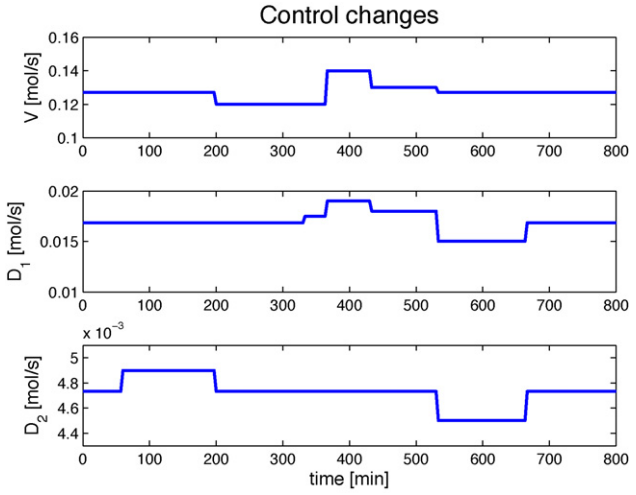


Fig. 2. Changes in $u = (D_1, D_2, V)$ to perturb the coupled column.

The measurement noise properties are assumed to be known, hence the covariance matrix $R \in \mathbb{R}^{53\times53}$ used in the estimators is $R = \mathrm{diag}(0.5^2, \ldots)$. The state noise covariance matrix $Q \in \mathbb{R}^{(106)\times(106)}$ and parameter covariance matrix $Q_p \in \mathbb{R}^{(3)\times(3)}$ are $Q = \mathrm{diag}(0.001^2, \ldots, 0.001^2)$ and $Q_p = \mathrm{diag}(0.001^2, 0.001^2, 0.001^2)$. These matrices are obtained by first making rough assumptions on a realistic magnitude of state noise and then modifying the matrices to achieve a satisfying EKF performance in the simulations. The very same covariance matrices are then used in the MHE formulation as weights, too. Finally, we set $Q_0^x = Q$ and $Q_0^p = Q_p$.

4.1.3. Discussion of estimation results

The process dynamics are excited by control changes and parameter changes shown in Figs. 2 and 3. While the controls are known to the estimators, the changing inflow parameters are not measured and have to be estimated. As an additional challenge for the estimators, inexact initial parameters and states are provided: Initial parameter as well as initial state estimates are set 20% lower than the true states. Note that this leads to inconsistent differential and algebraic states. While the multiple-shooting-based MHE algorithm can generally cope with algebraic inconsistencies, the EKF algorithm failed to recover from this inconsistency and crashed. Therefore, the perturbation of the algebraic initial guess had to be

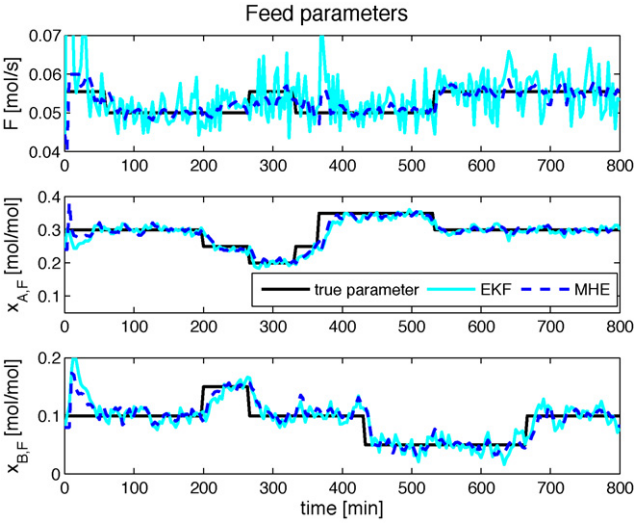


Fig. 3. True and estimated feed parameters $p = (F, x_{A,F}, x_{B,F})$.

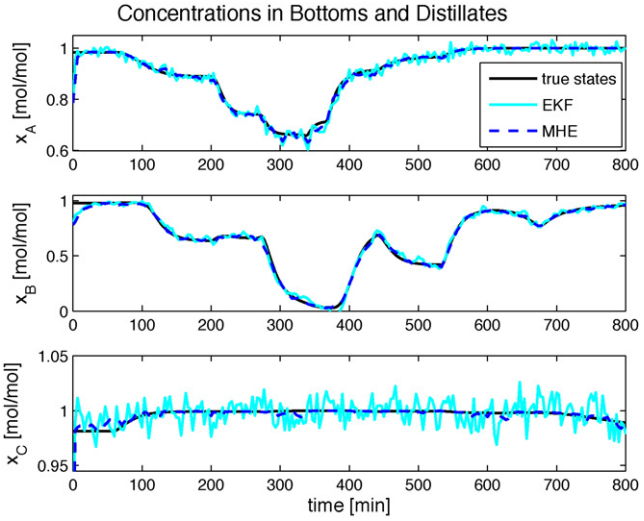


Fig. 4. True and estimated values for three concentrations (head 1, head 2 and bottoms).

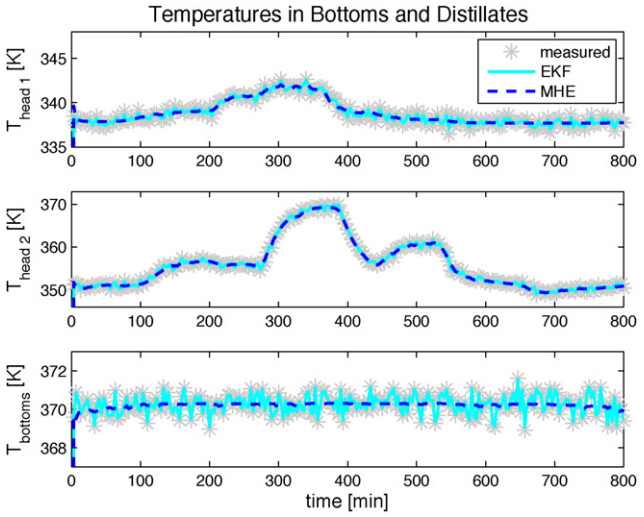


Fig. 5. Measured and estimated (i.e. filtered) values for three of ten measured temperatures (head 1, head 2 and bottoms).

reduced to 10% for the EKF to obtain meaningful estimates. In contrast to this, we were able to perturb the initial guesses of states and parameters by up to 60% before the MHE algorithm failed to provide meaningful results.

Selected product concentration estimates are shown in Fig. 4 together with their true values. The corresponding measured temperatures are displayed in Fig. 5.

It took an average CPU time per sampling interval of 4.00 s to solve the MHE optimization problem on a standard PC[4] and 4.29 s to compute the EKF estimates (see Fig. 6 for the CPU times). The MHE computation time is higher during the first number of samples because of additional background computations to obtain consistent states. In summary, the example demonstrates that the proposed MHE scheme reaches computation times of an order of magnitude that was only known for EKF algorithms before.

From the plots it seems that MHE is less sensitive to measurement noise and results in better estimates. This visual observation can be verified by comparing the accummulated relative error

[4] Intel Pentium 4 with 2.8 GHz, 1024 kB L2 cache, 1 GB RAM under Suse Linux 10.1.

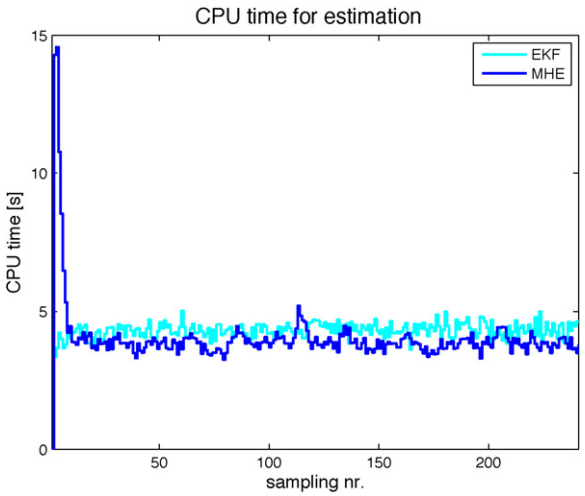


Fig. 6. Measured computation times of MHE and EKF for solving the estimation problem at a given sampling time.

(computed as sum of $\|\hat{x}_i - x_i\|_1 / \|x_i\|_1$ over all sampling instants i): it is 7.6 (MHE) and 12.9 (EKF) for the *estimated* states, and 0.3 (MHE and EKF) for the *filtered* (i.e. measured and estimated) states. The better MHE performance is due to the smoothing effect of taking a larger number of measurements into account as well as to the smoothing update used for the arrival cost. This effect has already been described by Haseltine and Rawlings (2005).

As for the bounds, a closer look at Fig. 4 shows that the EKF produces physically incorrect estimates for the three concentrations. They either exceed the value of 1.0 (x_A, x_C) or become slightly negative (x_B). This might be avoided by tedious tuning or by clipping values out of bounds. With MHE, these problems are avoided by explicitly taking constraints into account in the underlying nonlinear constrained optimization problem. The advanced handling of constraints can be seen as one of the major advantages of MHE as has already been pointed out by several researchers (e.g., Haseltine & Rawlings, 2005; Rao et al., 2003). This also holds for parameters as can be observed for the estimated feed flow F.

4.2. Application to the Tennessee–Eastman process

In the second example, both MHE algorithms are applied to the Tennessee Eastman (TE) benchmark process initially published in Downs and Vogel (1993). It has also served as an example for state estimation in Ricker and Lee (1995).

4.2.1. Model of the Tennessee–Eastman process

The TE process is a nonlinear, open-loop unstable process. It consists of five coupled unit operations: an exothermic two-phase continuous stirred-tank reactor, a vapor-liquid separator, a product condenser, a stripper and a recycle compressor. There are eight chemical components present in the process – components A, C, D and E are reactants, B is an inert contaminant, G and H are the desired products, and F is an unwanted byproduct. In the present work, we have used the extended model formulation described in Jockenhövel, Biegler, and Wächter (2003). Therein, the authors introduce an additional set of equations to describe the energy balances for the unit operations. At the same time, all base control structures have been removed, resulting in the open-loop instability. The flow sheet of the process is shown in Fig. 7.

The mathematical model consists of 30 differential and 140 algebraic equations and is based on a number of typical assumptions: All vapors behave as ideal gases, the vapor/liquid equilibrium follows Raoult's law, the vapor pressure is calculated using the Antoine equation, and all vessels are assumed perfectly mixed. Eight stream flow rates, two cooling water flow rates and one steam flow rate serve as controls.

4.2.2. MHE simulation study

Of the 170 process states, the following 30 are assumed to be measured: temperatures in the reactor, separator, and stripper and of the two cooling water flows; liquid levels of reactor, separator and stripper; reactor and separator pressure; the reactor feed rate and concentrations of the reactor feed stream, purge stream and final product stream. Additionally, the controls (i.e. control rates integrated over time) are assumed to be known. The measurements are available at each sampling instant with a sampling time of 100 s. For simplicity, all measurements come in with the same rate and without delay. Note, however, that the presented MHE algorithm can in principle cope with multi-rate measurements as well as delays as long as they are properly included into the least-

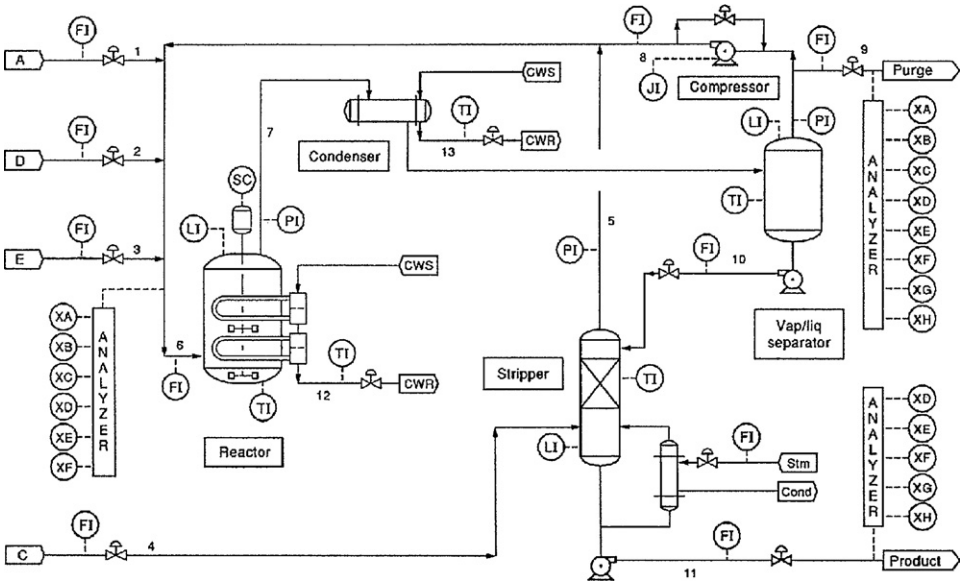


Fig. 7. Flow sheet of the Tennessee–Eastman Process (Jockenhövel et al., 2003).

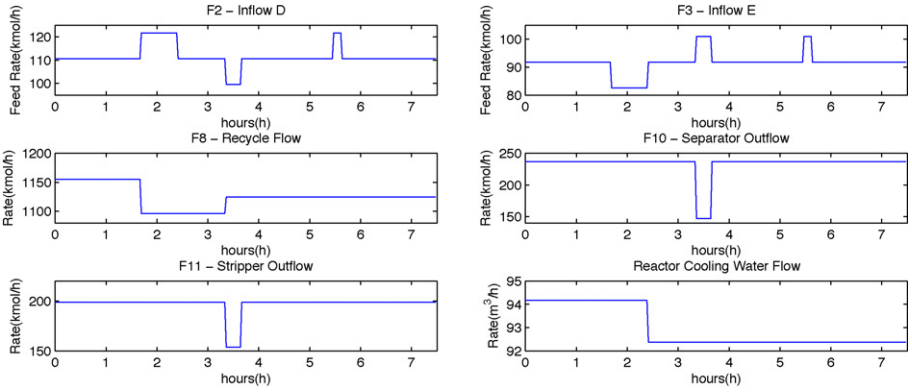


Fig. 8. Control profiles used for simulation of the Tennessee–Eastman process.

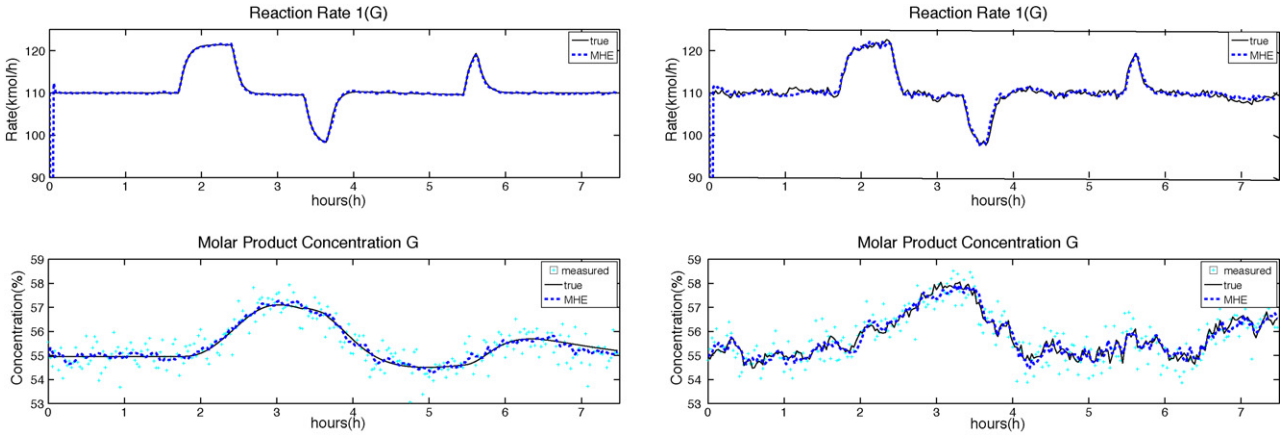


Fig. 9. (Upper left) Estimated and true state for output noise MHE. (Upper right) The same for output and state noise MHE. (Lower left) Filtered and true state along with measurements for output noise MHE. (Lower right) The same for output and state noise MHE.

squares problem. The estimation horizon is set to $T_E = 400$ s, which means that there are $M = 5$ sets of 30 temperature values available within each estimation window. Six of the process parameters are assumed unknown and have to be estimated along with the states. These parameters are: the concentration of reactants A and B in feed stream 4 (uncertain raw material supply), two pre-factors in the equations describing the reaction rates, and two pre-factors in the equations describing the liquid-vapor-equilibrium. The latter four parameters reflect uncertainties in the reaction kinetics and thermodynamics in the reactor.

Two different cases are treated in the simulations:

In the first case, we apply MHE formulation (3) *without* state noise terms. Note, however, that a certain degree of state noise is implicitly considered by the arrival cost update formulae.
In the second scenario, the differential states of the simulated process are indeed perturbed by Gaussian state noise. This permanently introduces small fluctuations that are not part of the estimator model. Standard deviation of the state noise is 0.5% of the respective initial values. The four temperatures, however, exhibit a higher deviation of 1% of the initial values. For this case, MHE formulation (4) is applied. In both scenarios, all measurements are corrupted by white Gaussian noise with standard deviations suggested in and.

The following covariance matrices are used for both scenarios: $R \in \mathbb{R}^{30\times30}$ is a diagonal matrix with the respective measurement noise variance as entries. Covariance matrix $Q \in \mathbb{R}^{30\times30}$ is a diagonal matrix of the state noise variance. For the parameter covariance matrix we have $Q_p \in \mathbb{R}^{6\times6}$ with $Q_p = \mathrm{diag}(10^{-5}, 10^{-5}, 10^{-4}, \ldots, 10^{-4})$.

The initial weight is based on $Q_0^x = 100\,Q$ and $Q_0^p = 100\,Q_p$. The simulation shows a change in product composition caused by a 10% feed change of components D and E. Some of the other flows are manipulated, too, to prevent violation of shutdown limits. The controls are shown in Fig. 8. Again, all calculations have been performed on a standard PC[5]. Integration accuracy is 10^{-5}.

4.2.3. *Discussion of estimation results*

An example of estimated and filtered states is displayed in Fig. 9 for both scenarios. The MHE algorithms capture the true state quickly despite wrong initial guesses. Results for the simultaneous parameter estimation are shown in Fig. 10 for output noise MHE and in Fig. 11 for the output and state noise variant. MHE is able to correctly estimate the parameter values despite sudden step changes. This task obviously becomes more difficult in face of the considerable level of state noise applied in the second scenario. The EKF results are not plotted in the same graphs for clarity of presentation.

The performance of EKF and MHE can be compared in Fig. 12. Note that a numerically more stable square-root variant of the EKF has been used here which is based on QR factorizations (Park & Kailath, 1995). The comparison shown is in relative errors $\|\hat{x}_i - x_i\|_1 / \|x_i\|_1$ at each sampling instant. The sum over all instants amounts to 0.3142 (EKF) and 0.2233 (MHE) in the case without

[5] Intel Pentium 4 with 2.8 GHz, 1024 kB L2 cache, 1 GB RAM under Suse Linux 9.3.

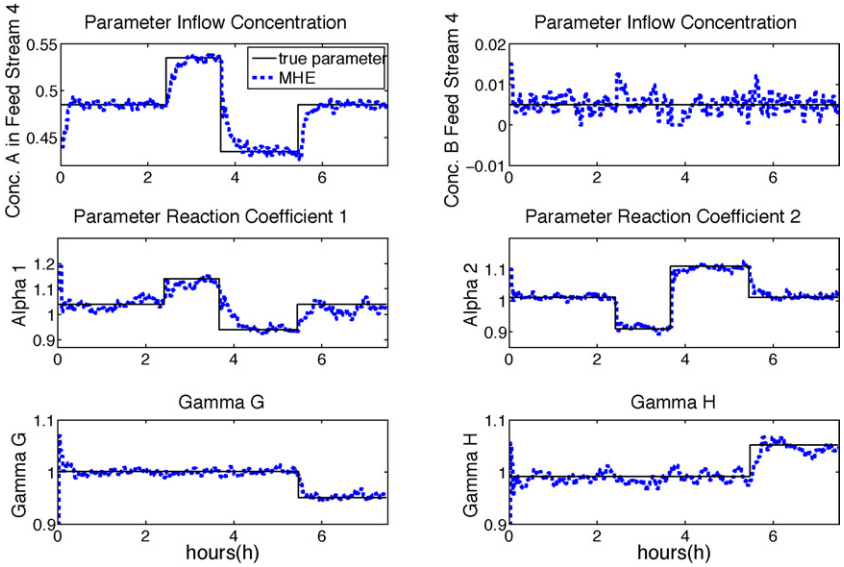


Fig. 10. Parameters estimated by output noise MHE.

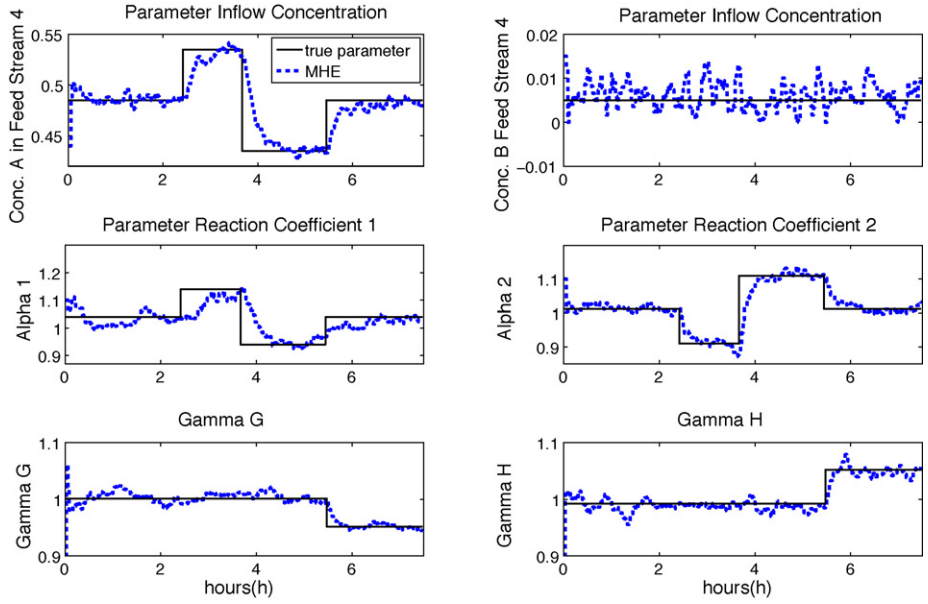


Fig. 11. Parameters estimated by output and state noise MHE.

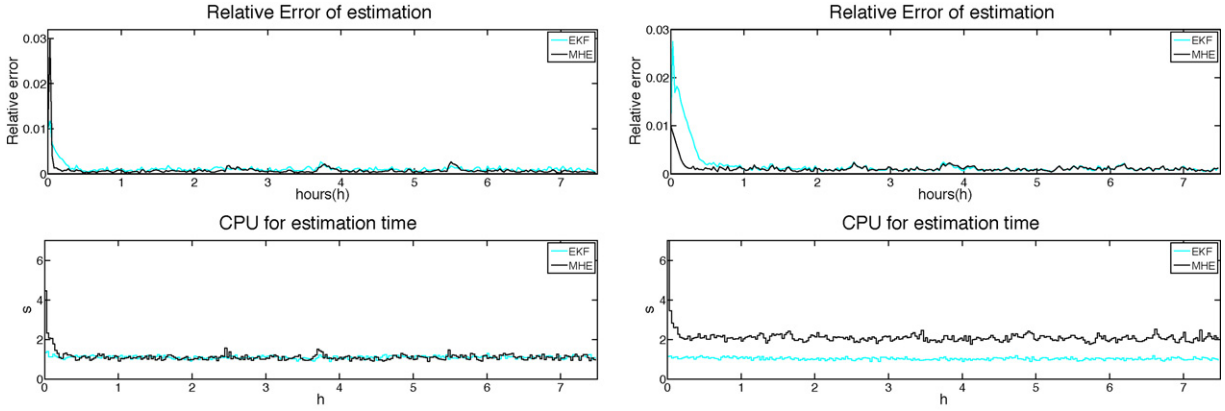


Fig. 12. A comparison of the general MHE (black/dark line) with results from EKF (cyan/bright line) by relative error $\|\hat{x}_i - x_i\|_1 / \|x_i\|_1$ of all states and by CPU time. (Left) Without state noise. (Right) With state noise.

state noise, and 0.4656 (EKF) and 0.3142 (MHE) in the case with state noise.

In both scenarios, the MHE is able to recover from bad initialization much quicker than the EKF. In terms of the accumulated relative errors after the initial phase, MHE and EKF perform equally well with slight advantages for MHE.

As for CPU times, findings from the first example are confirmed. The MHE version without state noise is in the range of one second per sampling interval just as the EKF. Only in the initial phase, computation time is higher with about 4.5 s for the very first estimate. The reason is the same as mentioned in Section 4.1. CPU times increase for the MHE version with state noise to a range of 2.2 s per sampling interval with the first estimate requiring almost 7 s. This is the price for a higher number of degrees of freedom in the optimization problem introduced to cover state noise. The CPU times for the EKF remain practically unchanged since the problem formulation has not been changed.

5. Summary

A moving horizon estimation (MHE) formulation to simultaneously estimate differential and algebraic states and unknown model parameters is presented. A derivation of the well-known arrival cost is presented that leads to an update scheme based on one single QR-factorization. To deliver estimates in real-time, a special numerical scheme – the MHE real-time iteration – is presented. Due to the multiple shooting approach and a careful shift strategy, the computational burden is reduced to one single generalized Gauss–Newton iteration per sample. This iteration can be split up into a preparation and an estimation phase, the latter being in the range of milliseconds. Two simulation examples show that the proposed MHE variant delivers superior estimates while being almost as fast (example 2) or even faster (example 1) than an Extended Kalman Filter. A formal proof of stability of the combined MHE and optimizer dynamics is subject to ongoing research.

Acknowledgments

The authors thank L. Wirsching, who helped implementing the MHE algorithm. The research was supported by the co-operative project cluster (DFG-Paketantrag) “Optimization-based control of processing plants” (DFG BO 864/10), as well as by Research Council KUL: CoE EF/05/006 Optimization in Engineering Center (OPTEC), and by Belgian Federal Science Policy Office: IUAP P6/04 (Dynamical systems, control and optimization, 2007-2011). The first author particularly acknowledges support by the international graduate college IGK710 “Complex processes: modeling, simulation, optimization”.

References

Bauer, I., Bock, H., & Schlöder, J. (1999). *DAESOL—A BDF-code for the numerical solution of differential algebraic equations*. Internal report, IWR, SFB 359, Universität Heidelberg.

Becerra, V., Roberts, P., & Griffiths, G. (2001). Applying the extended Kalman Filter to systems described by nonlinear differential-algebraic equations. *Control Engineering Practice*, *9*, 267–281.

Bestle, D., & Zeitz, M. (1983). Canonical form observer design for nonlinear time-variable systems. *International Journal of Control*, *38*, 419–431.

Binder, T., Blank, L., Bock, H., Bulirsch, R., Dahmen, W., Diehl, M., et al. (2001). Introduction to model based optimization of chemical processes on moving horizons. In M. Grötschel, S. Krumke, & J. Rambau (Eds.), *Online optimization of large scale systems: State of the art* (pp. 295–340). Springer.

Bock, H. (1981). Numerical treatment of inverse problems in chemical reaction kinetics. In K. Ebert, P. Deuflhard, & W. Jäger (Eds.), *Modelling of chemical reaction systems. Vol. 18 of Springer series in chemical physics* (pp. 102–125). Heidelberg: Springer.

Bock, H. (1983). Recent advances in parameter identification techniques for ODE. In P. Deuflhard, & E. Hairer (Eds.), *Numerical treatment of inverse problems in differential and integral equations* (pp. 95–121). Boston: Birkhäuser.

Bock, H., Diehl, M., Leineweber, D., & Schlöder, J. (2000). A direct multiple shooting method for real-time optimization of nonlinear DAE processes. In F. Allgöwer, & A. Zheng (Eds.), *Nonlinear predictive control. Vol. 26 of progress in systems theory* (pp. 246–267). Basel/Boston/Berlin: Birkhäuser.

Bock, H., & Plitt, K. (1984). A multiple shooting algorithm for direct solution of optimal control problems. In *Proceedings 9th IFAC world congress Budapest* (pp. 243–247). Pergamon Press.

Cannon, M. (2004). Efficient nonlinear model predictive control algorithms. *Annual Reviews in Control*, *28*, 229–237.

Daum, F. (2005). Nonlinear filters: Beyond the kalman filter. *IEEE A & E Systems Magazine*, *20*(8), 57–69.

Diehl, M., Bock, H., Schlöder, J., Findeisen, R., Nagy, Z., & Allgöwer, F. (2002). Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations. *Journal of Process Control*, *12*(4), 577–585.

Diehl, M., Findeisen, R., Allgöwer, F., Bock, H., & Schlöder, J. (2005). Nominal stability of the real-time iteration scheme for nonlinear model predictive control. *IEE Proceedings-Control Theory and Applications*, *152*(3), 296–308.

Diehl, M., Kuehl, P., Bock, H., & Schlöder, J. (2006). Schnelle Algorithmen für die Zustands- und Parameterschätzung auf bewegten Horizonten. *Automatisierungstechnik*, *54*(12), 602–613.

Downs, J., & Vogel, E. (1993). A plant-wide industrial process control problem. *Computers and Chemical Engineering*, *17*, 245–255.

Faanes, A., & Skogestad, S. (2005). Offset-free tracking of model predictive control with model mismatch: Experimental results. *Industrial and Engineering Chemistry Research*, *44*, 3966–3972.

Ferreau, H., Lorini, G., & Diehl, M. (2006). Fast nonlinear model predictive control of gasoline engines. In *Proceedings of the IEEE international conference on control applications* Munich, (pp. 2754–2759).

Findeisen, R., Allgöwer, F., & Biegler, L. (Eds.). (2006). *Assessment and future directions of nonlinear model predictive control. Lecture notes in control and information sciences*. Springer.

Gelb, A. (1974). *Applied optimal estimation*. Cambridge, MA: MIT Press.

Haseltine, E., & Rawlings, J. (2005). Critical evaluation of extended Kalman filtering and Moving-Horizon Estimation. *Industrial and Engineering Chemistry Research*, *44*, 2451–2460.

Itigin, A., Raisch, J., Moor, T., & Kienle, A. (2003, September 1–4). A two-level hybrid control strategy for the start-up of a coupled distillation plant. In *European control conference* Cambridge, UK.

Jockenhövel, T., Biegler, L. T., & Wächter, A. (2003). Dynamic optimization of the tennessee eastman process using the optcontrolcentre. *Computers and Chemical Engineering*, *27*, 1513–1531.

Jorgensen, J., Rawlings, J., & Jorgensen, S. (2004). Numerical methods for large-scale moving horizon estimation and control. In *Proceedings of international symposium on dynamics and control process systems (DYCOPS)*.

Kang, W. (2006). Moving horizon numerical observers of nonlinear control systems. *IEEE Transactions on Automatic Control*, *51*(2), 344–350.

Leineweber, D., Bauer, I., Schäfer, A., Bock, H., & Schlöder, J. (2003). An efficient multiple shooting based reduced SQP strategy for large-scale dynamic process optimization (Parts I and II). *Computers and Chemical Engineering*, *27*, 157–174.

Martinsen, F., Biegler, L. T., & Foss, B. A. (2004). A new optimization algorithm with application to nonlinear MPC. *Journal of Process Control*, *14*, 853–865.

M'hamdi, A., Helbig, A., Abel, O., & Marquardt, W. (1996). Newton-type receding horizon control and state estimation. In *Proceedings of the 13rd IFAC world congress* San Francisco, (pp. 121–126).

Michalska, H., & Mayne, D. Q. (1995). Moving horizon observers and observer-based control. *IEEE Transactions on Automatic Control*, *40*(6), 995–1006.

Muske, K., & Badgwell, T. (2002). Disturbance modeling for offset-free linear model predictive control. *Journal of Process Control*, *12*, 617–632.

Muske, K., & Edgar, T. (1997). Nonlinear state estimation. In M. Henson, & D. Seborg (Eds.), *Nonlinear process control*. Prentice Hall.

Muske, K., Rawlings, J., & Lee, J. (1993). Receding horizon recursive state estimation. In *Proceedings of the American control conference* San Francisco, (pp. 900–904).

Ohtsuka, T., & Fujii, H. (1996). Nonlinear receding-horizon state estimation by real-time optimization technique. *Journal of Guidance, Control, and Dynamics*, *19*(4).

Pannocchia, G., & Rawlings, J. (2003). Disturbance models for offset-free model-predictive control. *AIChE Journal*, *49*, 426–437.

Park, P., & Kailath, T. (1995). New square-root algorithms for Kalman filtering. *IEEE Transactions on Automatic Control*, *40*(5), 895–899.

Rao, C., Rawlings, J., & Lee, J. (2001). Constrained linear state estimation—A moving horizon approach. *Automatica*, *37*(2), 1619–1628.

Rao, C. V., Rawlings, J. B., & Mayne, D. Q. (2003). Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE Transactions on Automatic Control*, *48*(2), 246–258.

Rawlings, J., & Bakshi, B. (2006). Particle filtering and moving horizon estimation. *Computers and Chemical Engineering*, *30*, 1529–1541.

Ricker, N. L., & Lee, J. H. (1995). Nonlinear modeling and state estimation for the Tennessee Eastman challenge process. *Computers and Chemical Engineering*, *19*(9), 983–1005.

Robertson, D., & Lee, J. (1995). A least squares formulation for state estimation. *Journal of Process Control*, *5*(4), 291–299.

Robertson, D., Lee, J., & Rawlings, J. (1996). A moving horizon-based approach for least-squares state estimation. *AIChE Journal*, *42*, 2209–2224.

Schäfer, A., Kuehl, P., Diehl, M., Schlöder, J., & Bock, H. (2007). Fast reduced multiple shooting methods for nonlinear model predictive control. *Chemical Engineering and Processing*, *46*(11), 1200–1214.

Tenny, M. (2002, June). *Computational strategies for nonlinear model predictive control*. Ph.D. thesis, University of Wisconsin-Madison.

Tenny, M., & Rawlings, J. (2002). Efficient moving horizon estimation and nonlinear model predictive control. In *Proceedings of the American control conference* Anchorage, AK.

Tournambe, A. (1992). High-gain observers for nonlinear systems. *International Journal of Systems Science*, *23*, 1475–1489.

Ungarala, S. (2009). Computing arrival cost parameters in moving horizon estimation using sampling based filters. *Journal of Process Control*, *19*, 1576–1588.

Valdes-Gonzalez, H., Flaus, J.-M., & Acuna, G. (2003). Moving horizon state estimation with global convergence using interval techniques: Application to biotechnological processes. *Journal of Process Control*, *13*, 325–336.

Zavala, V., Laird, C., & Biegler, L. (2006). Fast solvers and rigorous models: Can both be accommodated in NMPC? In *Proceedings of the IFAC workshop on nonlinear model predictive control for fast systems* Grenoble.

Zavala, V., Laird, C., & Biegler, L. (2007). A fast computational framework for large-scale moving-horizon estimation. In *Proceedings of the 8th international symposium on dynamics and control of process systems*.