Moving Horizon Estimation - The Explicit Solution

John D. Hedengren and Thomas F. Edgar*
Department of Chemical Engineering
University of Texas at Austin
Austin, TX 78712

April 11, 2005

Abstract

Moving horizon estimation consistently outperforms the Kalman filter with greater robustness to both poor initial state guesses and sub-optimal estimator tuning parameters. The only price of improvement is the greater computational expense required to solve the MHE optimization. The contribution in this work is that the tradeoff of greater computational expense is eliminated by developing an explicit solution to the MHE optimization problem. Unlike the implicit optimization approach, the explicit solution result is guaranteed in a highly predictable computational time that is minimal even for large-scale nonlinear models with long time horizons. For state estimation problems with inequality constraints, an iterative layer is added to determine the set of active constraints. An augmented objective function monitors the solution progression to guarantee convergence.

Keywords

Moving Horizon Estimation, Kalman Filter, State and Parameter Estimation

Introduction

Estimation of a model states and parameters from process measurements bridges the gap between the theoretical realm of mathematical models and the realistic processes they represent. Many approaches have been suggested to accomplish the reconciliation of model and process, with a range of tradeoffs (Soroush, 1998). Generally, the tradeoffs are centered on model form and size, computational expense, ease of implementation, robustness to process/model mismatch, and cultural factors such as understanding and acceptance.

The reconciliation process is an important precursor to many other activities such as fault detection, product quality assurance, manual control, and model-based control. These model-based techniques need an accurate estimate of the current system variables to perform well. Without accurate state estimation, many of these tools would perform poorly or fail.

Previous Work

For dynamic nonlinear model-based control of chemical processes, the most popular feedback strategies in practice are the extended Kalman filter and a constant or integrating output disturbance variable (Qin and Badgwell, 2000). The Kalman filter is optimal for unconstrained, linear systems subject to known normally distributed state and measurement noise (Haseltine and Rawlings, 2004). The Kalman filter sequentially updates state estimates based on the magnitude of the error between the measurements and the model variables. The Kalman filter is simply an optimal proportional-only filter that proportionally corrects state values from the deviation of model values from measurements. The extended Kalman filter is an extension of the Kalman filter, developed for unconstrained, nonlinear DAE systems (Becerra et al., 2001). By linearizing the model about updated state estimates, the extended Kalman filter is able to predict the nonlinear state evolution, although sub-optimally (Haseltine and Rawlings, 2004). Vachhani et al. (2005) proposed EKF with constraints, although the augmentation strategy for parameter estimation is still a limitation.

A number of critical evaluations have shown that moving horizon estimation (MHE) consistently outperforms the extended Kalman filter (Haseltine and Rawlings, 2004) (Jang et al., 1986) (Robertson and Lee, 1995). State estimation of real systems may include changing measurement frequencies, multiple measurements at different sampling frequencies, measurement delay, large-scale nonlinear models, and constraints. MHE is

*edgar@che.utexas.edu

an optimization-based approach that is flexible to deal with these complications (Michalska and Mayne, 1995) (Moraal and Grizzle, 1995). MHE is also known as nonlinear dynamic data reconciliation (NLDDR) (Liebman et al., 1992) (Soderstrom et al., 2000). MHE is a computationally tractable approximation to the optimal infinite horizon estimation (Rao et al., 2001). All of the challenges of real system state estimation are naturally handled in the MHE framework. An estimate of the current states is typically obtained by solving a least squares optimization problem subject to the model constraints and inequality constraints that represent bounds on variables or equations. Most of the published work centers around different techniques that solve the same minimization problem. Jang et al. (1986) iteratively linearized the nonlinear ODE model about a reference trajectory by computing sensitivities. Liebman et al. (1992) first proposed a simultaneous NLP solution approach where the differential equations are transformed into algebraic equations through orthogonal collocation on finite elements. Ramamurthi et al. (1993) proposed a two step process to implicitly estimate the input disturbances while explicitly calculating state estimates. Albuquerque and Biegler (1995) exploited the MHE SQP structure to achieve linear computational scaling with horizon length for ODE models. They later extended the technique to DAE systems (Albuquerque and Biegler, 1997).

A number of enhancements have extended the theoretical basis and functionality of MHE. (McBrayer and Edgar, 1995) proposed a bias detection and estimation strategy to improve state estimation. Offset free estimation and control is achieved by augmenting the model with a number of disturbance variables equal to the number of measurements (Muske and Badgwell, 2002) (Pannocchia, 2003). Rao et al. (2001) derived sufficient conditions for MHE with linear systems subject to constraints. They also suggested an infinite horizon approximation by weighting previous state estimates in the least squares problem.

Moving Horizon Estimation Problem Formulation

The MHE optimization is typically a minimization of a least squares objective function to allign the model with measured values.

$$\begin{aligned} \min J &= \|Y_s - Y_m\|_{Q_y}^2 \\ \text{s.t.} \quad 0 &= f(\dot{x}, x, u, p) \\ y_s &= g(x, u, p) \\ a &\geq h(x, p) \geq b \end{aligned} \tag{1}$$

where J is the objective function value, Y_s is a vector of measurements at all nodes, Y_m is a vector of model values at the sampling times, Q_y is the inverse of the measurement covariance, f is a vector of model equation residuals, x represents the model states, u is the model inputs, p is the model parameters, y_s is a vector of measurements, g is an output function, h is an inequality constraint function, and a and b are lower and upper limits, respectively. Sensitivities of the initial conditions are computed to discretize the nonlinear model. In practice, this discretization step is the most computationally expensive part of the MHE calculation. For this study, it is assumed that the discrete model is freely available. The vectors y_m and y_s are successively stacked to form Y_m and Y_s where the horizon length is n.

$$Y_m = \begin{bmatrix} y_{m,0} \\ \vdots \\ y_{m,n} \end{bmatrix}, \quad Y_s = \begin{bmatrix} y_{s,0} \\ \vdots \\ y_{s,n} \end{bmatrix} \tag{2}$$

An infinite horizon approximation is added by incorporating a penalty on the deviation from previous model estimates. This penalty is added by augmenting the objective function with the least squares contribution of previous model estimates $\hat{X}_m$, weighted with a forgetting factor α. Disturbance variables (shown here as input disturbances), d, are included as state variables to achieve offset free estimation and control. The nonlinear inequality constraints are simplified by defining new states $z_k = h(x_k, p_k)$ and imposing inequality constraints on z_k.

$$\begin{aligned} &\min J = \|X_s - X_m\|_{Q_x}^2 + \alpha \left\|\hat{X}_m - X_m\right\|^2 \\ &\text{s.t.} \begin{bmatrix} x \\ d \\ p \end{bmatrix}_{k+1} = \begin{bmatrix} A & B & P \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}_k \begin{bmatrix} x \\ d \\ p \end{bmatrix}_k + \begin{bmatrix} B \\ 0 \\ 0 \end{bmatrix}_k u_k \\ &\quad y_{s,k} = \begin{bmatrix} C & 0 & 0 \end{bmatrix}_k \begin{bmatrix} x \\ d \\ p \end{bmatrix}_k \\ &\quad a \geq z_k \geq b \end{aligned} \tag{3}$$

The vectors $\hat{X}_m$ and X_m are stacked model vectors $\hat{x}_m$ and x_m. Also, X_s is constructed from the measurements and $Q_{x,k} = \left(C_k^T Q_{y,k} C_k\right)$.

$$X_m = \begin{bmatrix} x_{m,0} \\ \vdots \\ x_{m,n} \end{bmatrix}, \quad \hat{X}_m = \begin{bmatrix} \hat{x}_0 \\ \vdots \\ \hat{x}_n \end{bmatrix} \tag{4}$$

$$X_s = \begin{bmatrix} \left(C_0^T C_0\right)^{-} C_0^T \left(y_{s,0} - D_0 u_0\right) \\ \vdots \\ \left(C_n^T C_n\right)^{-} C_n^T \left(y_{s,n} - D_n u_n\right) \end{bmatrix} \tag{5}$$

Solution of this optimization problem is typically accomplished with an implicit solution technique. The next section develops an explicit solution to this problem.

The Explicit MHE Solution

For simplicity of the derivation, the augmented state matrix is reduced to a generic linear time-varying form of $x_{k+1} = A_k x_k + B_k u_k$ and $y_k = C_k x_k + D_k u_k$. All variables are in deviation form although not explicitly indicated. The model evolution is a unique function of the initial states.

$$\begin{aligned}
&\omega_k = \prod_{j=0}^{k-1} A_j \quad \psi_k = \sum_{j=1}^{k} \left[\prod_{i=1}^{j-1} A_{i-k-j}\right] B_{k-j} u_{k-j} \\
&\Omega = \begin{bmatrix} I \\ \omega_1 \\ \vdots \\ \omega_{n-1} \end{bmatrix} \quad \Psi = \begin{bmatrix} 0 \\ \psi_1 \\ \vdots \\ \psi_{n-1} \end{bmatrix} \\
&X_m = \Omega x_0 + \Psi
\end{aligned} \tag{6}$$

The equations of X_m and Y_m are substituted into the objective function making it a unique function of x_0. The explicit solution to the minimization problem is obtained by differentiating the objective function with respect to x_0, setting the differentiated function equal to zero, and algebraically manipulating the equation to solve for the estimated x_0 ($\hat{x}_0$) explicitly.

$$\begin{aligned}
&\hat{x}_0 = T^{-1} \left(\Omega^T \left(Q_x X_s + \alpha X_m - (Q_x + \alpha I)\, \Psi\right)\right) \\
&\text{with } T = \Omega^T \left(Q_x + \alpha I\right) \Omega
\end{aligned} \tag{7}$$

The explicit solution can be calculated when the inverse of T exists. T^{-1} exists when previous estimates are used to approximate the infinite horizon solution ($\alpha > 0$). T is not invertible when the system is unobservable and $\alpha = 0$. This property is consistent with the fact that an unobservable system possesses extra degrees of freedom leading to states that cannot be estimated from the available measurements.

Inequality Constraints in Explicit MHE

Inequality constraints represent physical limits on state variables or combinations of state variables. For example, mole fractions are always between 0 and 1. If the state estimation predicted a mole fraction outside of this range, that mole fraction would have little physical meaning and would decrease the credibility of the other results. Inequality constraints add valuable information to the state estimation. For systems that are partially unobservable, the inequality constraints bound the unobservable states, thereby increasing the level of system observability. However, an unobservable system cannot be made completely observable with inequality constraints. Additional actual measurements are the only way to make an unobservable system completely observable.

As previously mentioned, the inequality constraints $a \geq h(x_k, p_k) \geq b$ are simplified by creating new variables z_k and adding $z_k = h(x_k, p_k)$ to the set of state equations. Equivalent constraint information is retained by imposing inequality constraints on z_k ($a \geq z_k \geq b$). Imposing constraint information leads to a possible infeasible solution. To overcome this possibility, the inequality constraints are ranked according to the order of importance. This ranking is accomplished by softening the constraints and imposing successively higher weighting on more important constraints. Softening the constraints guarantees a feasible solution because the inequality constraints may be violated to meet the state equality constraints. Softening of the constraints is performed in practice by adding a penalty to the objective function for constraint violation.

$$\begin{aligned}
&\min J + s_a^T Q_a s_a + s_b^T Q_b s_b \\
&\text{s.t.} \quad \text{the state equations} \\
&\qquad s_a = a - X_m \\
&\qquad s_b = X_m - b
\end{aligned} \tag{8}$$

The matrices Q_a and Q_b have diagonal elements that turn on (weighting > 0) or off (weighting $= 0$) to control the set of active constraints. A MHE problem with inequality constraints is iterative because the final set of active constraints is not known a priori. However, the prediction of states, disturbances, and parameters is still an explicit solution for a known set of active inequality constraints. The computational time required to solve a problem with inequality constraints is variable, equal to the time required for one explicit solution multiplied by the number of iterations. The explicit solution given a set of active inequality constraints is given in Equation 9.

$$\begin{aligned}
&\hat{x}_0 = T^{-1} \left(\Omega^T \left(Q_x X_s + \alpha X_m + Q_a a + Q_b b - R\Psi\right)\right) \\
&\text{with } R = \left(Q_x + \alpha I + Q_a + Q_b\right) \text{ and } T = \left(\Omega^T R \Omega\right)
\end{aligned} \tag{9}$$

Convergence is guaranteed by rejecting iterations that give do not produce a sufficient decrease in the objective function. There are many strategies to accomplish guaranteed convergence, although some strategies prove superior to others. Each iteration is a new set of active constraints that are predicted to give a decrease in the objective function. The initial set of active constraints is determined by computing an unconstrained MHE solution. Weighting is added to Q_a and Q_b for states that violate the inequality constraints. In successive iterations, weighting is removed for constraints with negative Lagrange multipliers ($\lambda_a = -2Q_a s_a$ and $\lambda_b = 2Q_b s_b$. If the new set of active constraints does not give a sufficient decrease in the objective function, the algorithm adjusts the weights on the active constraints. The parameter β is decreased until a decrease in the objective function is discovered.

$$\begin{aligned}
Q_a &= \beta Q_{a,k} + (1-\beta) Q_{a,k-1} \\
Q_b &= \beta Q_{b,k} + (1-\beta) Q_{b,k-1}
\end{aligned} \tag{10}$$

Once a decrease in the objective function is discovered, the optimizer accepts the iteration and moves on to find a new search direction (new active set of constraints that is predicted to produce a decrease of the objective function). This iteration sequence is terminated when the active set does not change from one iteration to the next.

Example 1: Flash Column Composition Estimation

This section shows an example of MHE, but for a physically realistic process. The third example is a 17 state model of a flash column. The unknown compositions are estimated from the temperature and flow rate measurements. A flash column is an elementary separation unit that is fed with a pressurized liquid stream. When the liquid stream enters the column at a reduced pressure that is below the liquid's vapor pressure, a fraction of the stream instantaneously flashes into the vapor phase. A rigorous nonlinear model of mass, energy, and thermodynamic equilibrium relationships predicts the dynamic behavior of the column. A diagram of the model is presented in Figure 1. The feed tank contains an equimolar hydrocarbon mixture of C_4H_{10}, C_5H_{12}, C_6H_{14}, C_7H_{16}, and C_8H_{18}. The feed and flash temperatures are measured as are the vapor and liquid flow rates. Noise is added to the measurements with mean zero and standard deviation 0.5 for the temperatures and 0.02 for the flow rates. The objective is to estimate the feed tank compositions from the temperature and flow rate measurements. Figure 2 shows the measurements taken over the time horizon of interest. The 17 state model has 5 differential states and 12 algebraic states. For full observability, the observability matrix must have rank ≥ 5. For this example, the observability matrix is rank deficient at 3. This analysis reveals that temperature and flow measurements of a flash column can only be used to exactly estimate compositions of mixtures with ≤ 3 components. Alternatively, 2 additional compositions could be measured to make the system observable. However, even if the system is not fully observable, some information can be reconstructed that can be better than the initial composition estimates. For this example, the compositions are initially estimated as 0.3 whereas the actual compositions are all located at 0.2. Figure 3 shows the estimation of the compositions over a 100 second time horizon. A forgetting factor of 0.5 on the initial state was used to incorporate previous estimates. The estimation is able to reconstruct the compositions of C_4H_{10}, C_5H_{12}, and C_6H_{14}. However, the other two composi-

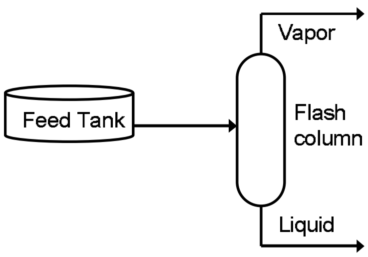

Figure 1: Flow sheet diagram of the flash column model. The flash column consists of a feed tank with unknown species compositions, a flash column, and vapor and liquid outlet streams.

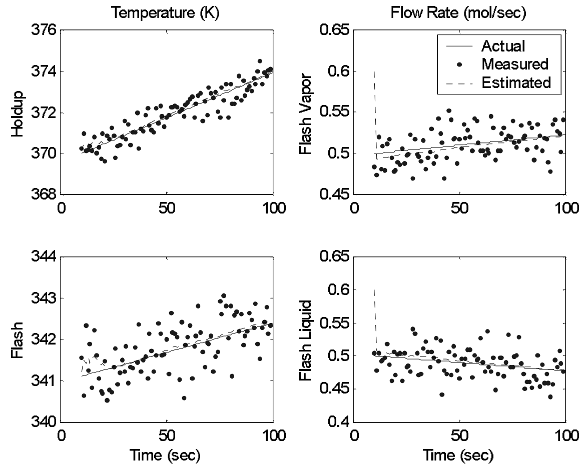

Figure 2: The estimated states converge quickly to the real system even though the initial guess is poor.

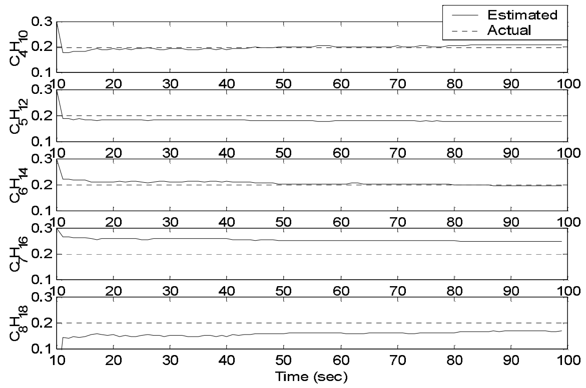

Figure 3: Estimated compositions of C_4H_{10}, C_5H_{12}, and C_6H_{14} approach the actual values of 0.2. The other two compositions of C_7H_{16} and C_8H_{18} deviate significantly because the system is not fully observable.

tions, C_7H_{16} and C_8H_{18}, deviate significantly from the correct solution. This deviation is a result of an unobservable system.

Inequality constraints can bound unobservable states to increase the accuracy of the estimation. For this example, suppose it is known that the composition of C_7H_{16} should not be above a composition of 0.22. This constraint information can be incorporated into the explicit MHE formulation to provide a better estimate of compositions. Figure 4 shows the results of bounding the C_7H_{16} composition. At the final solution the active

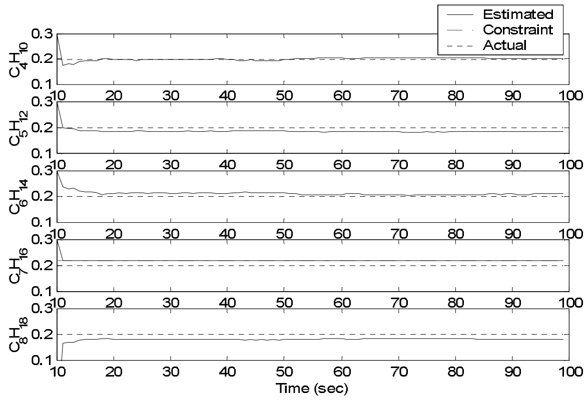

Figure 4: The composition estimation is greatly improved by adding an inequality constraint to C_7H_{16}. Even though the system is not fully observable, the compositions closely approximate the actual values.

constraint on C_7H_{16} composition has a Lagrange multiplier of +0.02, confirming that the constraint should be active. The estimation of the composition is greatly improved by incorporating additional information about the process in the form of an inequality constraint.

Explicit MHE Scaling with Model Size

An important property of explicit MHE is computational scaling to large-scale problems. To test the scalability to large-scale problems, a series of 17 state flash columns are combined to form larger models. These successively larger models are solved for the linear and nonlinear case as seen in Figure 5. A horizon of 50 samples is used for all of the simulations. Both linear and nonlinear explicit MHE scale $O(x^2)$ in the number of floating point operations, although the linear approach scales approximately 6 times better than the nonlinear method. With computers that operate in the Gigahertz range, the computational feasibility of explicit MHE is excellent even for large-scale problems (10,000+ variables).

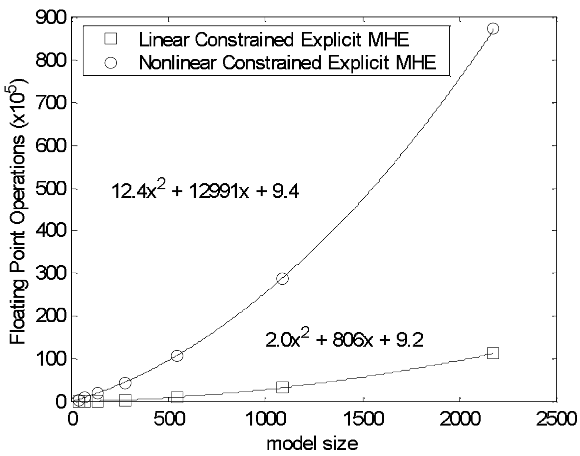

Figure 5: Explicit MHE scaling to large-scale model size. Both the nonlinear and linear approaches scale $O(x^2)$ in the number of floating point operations where x is the number of variables in the model.

Explicit MHE Scaling to Long Time Horizons

Some estimation problems require long time horizons ($>$ 100 sampling intervals). Long time horizons may be necessary when the measurements have low signal to noise ratios, process measurements occur much faster than the process dynamics, or there is a large difference among the sampling frequencies of multiple measurements. Another reason for a long time horizon is for parameter estimation where a few parameters are estimated from a long time period of historical data. Figure 6 displays the effect of time horizon length on the number of floating point operations for the 17 state flash column model. For nonlinear models, the scaling is quadratic for increasing horizon length. For linear models the scaling is linear for increasing horizon length. The linear model scaling is particularly amenable for problems that may require a very long time horizon.

Example 2: Two State CSTR

State estimation of a CSTR is a popular benchmark test problem as found in Albuquerque and Biegler (1995), Haseltine and Rawlings (2004), Jang et al. (1986), Liebman et al. (1992), McBrayer and Edgar (1995), Ramamurthi et al. (1993), and Rao and Rawlings (2002) . The purpose of this example is to estimate the computational load for different estimation strategies.

A realistic estimation problem was devised to test eMHE for a sequence of step responses. The estimator horizon is set to 60 minutes and divided into 1 minute segments. The temperature is sampled every minute and

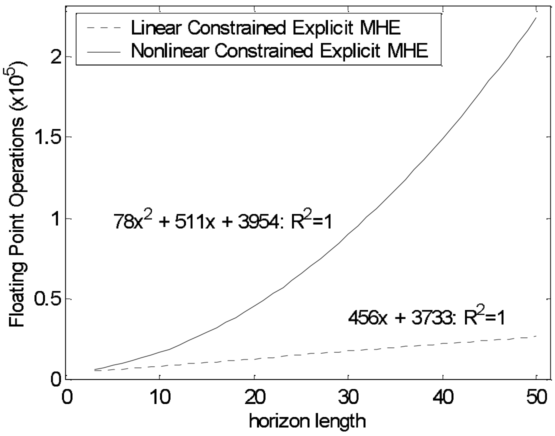

Figure 6: Explicit MHE scaling to horizon length. For nonlinear models, scaling is $O(x^2)$ in the number of floating point operations. For linear models, scaling is $O(x)$ where x is the horizon length.

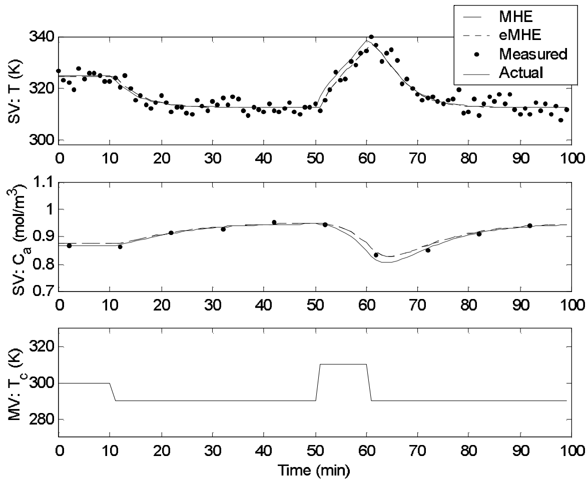

Figure 7: Estimation performance of the explicit solution MHE (eMHE) versus MHE. The state variable (SV) estimation is difficult to distinguish on the graph because the predictions are virtually identical for the two approaches. The only difference is the substantially lower computational effort of eMHE.

corrupted by normally distributed noise with a standard deviation of 5 K. Concentration is sampled every 10 minutes with a standard deviation of 0.01 $\frac{mol}{m^3}$. Plant-model mismatch is introduced by using an activation energy of the first order ($A \rightarrow B$) reaction of 8750 $\frac{J}{mol}$ for the model and 8740 $\frac{J}{mol}$ for the plant. The plant-model mismatch is introduced to cause deviation of the estimated response from the actual process. The steady state deviation can be eliminated by including parameter estimation or a disturbance variable. At the first sampling time the plant is assumed to be at steady state with a jacket cooling temperature of 300 K. At 20 minutes the cooling temperature is set to 290 K, followed by a step to 310 K at 60 minutes. At 70 minutes the cooling temperature returns to 290 K. Figure 7 shows the results of the MHE study. The eMHE solution averaged approximately 22,000 floating point operations to compute a solution. The direct single shooting optimization MHE solution averaged approximately 40 million floating point operations. The CPU time results from Liebman et al. were performed on a computer that delivers approximately 1 MFLOPS with LINPACK benchmark tests (Liebman et al., 1992). He reported in 1992 solution times in the range of 1-100 seconds giving approximate computational effort in the range of 1-100 million floating point operations for sparse solvers and orthogonal collocation on finite elements. The explicit solution approach offers improved computational performance that is insensitive to convergence tolerance, poor initial conditions, strong nonlinearities, and other factors that influence the implicit solution approach.

Conclusions

Moving horizon estimation has been established as a superior state estimation technique compared with the extended Kalman filter. The only established tradeoff is the additional computational expense need to solve the MHE optimization problem. An explicit solution removes the computational disadvantage for large scale nonlinear DAE systems that is guaranteed when the system is fully observable or when previous estimates are incorporated into the optimization. Inequality constraints add variable bounds that can improve the state estimation, especially for systems that are not fully observable. An iterative approach is necessary to determine an active set of equality constraints from the full set of inequality constraints. The iterative solution has guaranteed convergence by selecting new active sets that generate a sufficient decrease in the objective function. The computational expense of the most challenging problem in this paper required 22,000 floating point operations, only a few micro-seconds with modern computational power. The computational expense of implicit optimization MHE is significantly more, with a possibility of convergence failure depending on the initial conditions selected, problem nonlinearity, choice of optimizer, etc.

References

Albuquerque, J. and L. Biegler, “Decomposition Algorithms for On-Line Estimation with Nonlinear Models," *Computers and Chemical Engineering*, **19**(10), 1031–1039 (1995).

Albuquerque, J. and L. Biegler, "Decomposition Algorithms for On-Line Estimation with Nonlinear DAE Models," *Computers and Chemical Engineering*, **21**(3), 283–299 (1997).

Becerra, V., P. Roberts, and G. Griffiths, "Applying the Extended Kalman Filter to Systems Described by Nonlinear Differential-Algebraic Equations," *Control Engineering Practice*, **9**, 267–281 (2001).

Haseltine, E. and J. Rawlings, "Critical Evaluation of Extended Kalman Filtering and Moving-Horizon Estimation," *Ind. Eng. Chem. Res.* (2004).

Jang, S., B. Joseph, and H. Mukai, "Comparison of Two Approaches to On-Line Parameter and State Estimation of Nonlinear Systems," *Ind. Eng. Chem. Process Des. Dev.*, **25**, 809–814 (1986).

Liebman, M., T. Edgar, and L. Lasdon, "Efficient data reconciliation and estimation for dynamic processes using nonlinear programming techniques," *Computers and Chemical Engineering*, **16**, 963–986 (1992).

McBrayer, K. and T. Edgar, "Bias detection and estimation in dynamic data reconciliation," *Journal of Process Control*, **5**(4), 285–289 (1995).

Michalska, H. and D. Mayne, "Moving Horizon Observers and Observer-Based Control," *IEEE Transactions on Automatic Control*, **40**(6), 995–1006 (1995).

Moraal, P. and J. Grizzle, "Observer Design for Nonlinear Systems with Discrete-Time Measurements," *IEEE Transactions on Automatic Control*, **40**(3), 395–404 (1995).

Muske, K. R. and T. A. Badgwell, "Disturbance Modeling for Offset-Free Linear Model Predictive Control," *Journal of Process Control*, **12**, 617–632 (2002).

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

Qin, S. and T. Badgwell, *Nonlinear Model Predictive Control*, chapter An overview of nonlinear model predictive control applications, pages 369–392. Birkhäuser Verlag, Boston, MA (2000).

Ramamurthi, Y., P. Sistu, and B. Bequette, "Control-Relevant Dynamic Data Reconciliation and Parameter Estimation," *Computers and Chemical Engineering*, **17**(1), 41–59 (1993).

Rao, C. and J. Rawlings, "Contrained process monitoring: Moving-horizon approach," *AIChE Journal*, **48**(1), 97–109 (2002).

Rao, C., J. Rawlings, and J. Lee, "Constrained linear state estimation - a moving horizon approach," *Automatica*, **37**, 1619–1628 (2001).

Robertson, D. and J. Lee, "A Least Squares Formulation for State Estimation," *Journal of Process Control*, **5**(4), 291–299 (1995).

Soderstrom, T., T. Edgar, L. Russo, and R. Young, "Industrial Application of a Large-Scale Dynamic Data Reconciliation Strategy," *Industrial and Engineering Chemistry Research*, **39**, 1683–1693 (2000).

Soroush, M., "State and parameter estimations and their applications in process control," *Computers and Chemical Engineering*, **23**, 229–245 (1998).

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