Advanced-multi-step Moving Horizon Estimation

Yeonsoo Kim * **Kuan-Han Lin** ** **David M. Thierry** **
Lorenz T. Biegler **,[1]

* *Department of Chemical Engineering, Kwangwoon University, 20 Kwangwoon-ro, Nowon-gu, Seoul, 01897, Republic of Korea*
** *Chemical Engineering Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA*

Abstract: Moving Horizon Estimation (MHE) is an important optimization-based approach for state estimation and parameter updates, because of its capabilities in dealing with nonlinearity and state constraints. In addition, one of the applications is to provide the full state information for Model Predictive Controller (MPC) to control the process in either setpoint tracking or economic control purposes. However, the computational burden of MHE could deteriorate the control performance if the feedback delay caused by computation is too long, leading to potential safety issues or process damage. In this paper, we propose a fast moving horizon estimation algorithm to overcome the long computational time of MHE for real-time control applications, especially for fast dynamics or large-scale systems. We exploit the nonlinear programming (NLP) sensitivity and make use of efficient NLP solvers, IPOPT and *k_aug*, to reduce the on-line computational costs. This new approach is demonstrated on a CSTR process, where results are compared to ideal MHE and advanced-step MHE (asMHE).

Keywords: Estimation Algorithms, Optimal Estimation, Nonlinear Programming, Sensitivity Analysis, Real-time

1. INTRODUCTION

Moving Horizon Estimation has recently become more popular, especially for model based control strategies such as Nonlinear Model Predictive Control (NMPC). By solving an optimization problem at each sampling time that possesses a sequence of noised measurements, MHE provides the full state information and real-time updates for changing parameters. The former information can be set as the required initial condition of NMPC and the latter updates can improve the performance of NMPC with more accurate parameters. In addition, compared to the well-known Extended Kalman Filter (EKF) (Bryson and Ho (1975), Jazwinski (1970)), the advantages of MHE are its abilities to handle nonlinearities with first principles models and to deal with constraints in a simple and consistent way (Haseltine and Rawlings (2005)). These constraints enable us to include more physical information of the states, leading to better performance of the estimator by avoiding the infeasible region.

However, the non-negligible on-line computing time requried for MHE raises some concerns. Reducing computational delay is also important for chemical processes. A delayed state estimate incurs a deadtime in control action, which leads to loss of control performance as well as possible destabilization of the process (Findeisen and Allgöwer (2004), Diehl et al. (2005), Chen et al. (2000)). Some methods to cope with on-line computational load include applying the generalized Gauss-Newton method to a simultaneous framework of the estimation problem in Kraus et al. (2006) and Wynn et al. (2014). In particular, Zavala et al. (2008) proposed a real-time MHE algorithm to tackle the online computational burden with nonlinear programming (NLP) sensitivity. In his approach, the measurements for the next step are predicted by the plant model and the MHE problem is solved *offline* with the predicted measurements in background. Once the actual measurements are obtained, the optimal solution based on the true measurements is updated quickly using NLP sensitivity. This online updating approach only performs a backsolve for the linear system in each step, and requires less computation time compared to solving a full MHE problem. Nonetheless, this approach can only apply to cases where the MHE solution time is less than one sampling time; this may not be suitable for fast dynamic systems or large-scale plants.

Following on our previous work on advanced-multi step NMPC (amsNMPC, Kim et al. (2020b)), we propose the advanced-multi step MHE (amsMHE) requiring only one processor to handle the longer solution time for MHE. Basically, we predict a number of future measurements with plant model and solve the MHE a few steps ahead depending on how long it takes to obtain the optimal solution in background. When we obtain the actual measurements, we place them in the corresponding position inside the horizon window, update the solution with NPL sensitivity, and retrieve the current estimates. The main difference from asMHE is that we now solve the MHE problem over

[1] Corresponding author. Tel.: +1-412-268-2232; Fax: +1-412-268-7139; E-mail address: lb01@andrew.cmu.edu

multiple sampling times. In addition, the results from this MHE problem would then be used for multiple time steps until the next MHE result is available.

This paper is organized as follows. In section 2, we introduce the formulation of MHE and NLP sensitivity. Section 3 presents the algorithm of amsMHE. Section 4 shows the CSTR case study results and compares them to ideal MHE and asMHE. Finally, section 5 concludes this paper and provides some future perspectives.

2. IDEAL MHE AND NLP SENSITIVITY

Consider the nonlinear system having the state disturbances w_k and measurement noise v_k:

$$\begin{aligned} x_{k+1} &= f(x_k, w_k) \\ y_k &= h(x_k) + v_k. \end{aligned} \tag{1}$$

At time step t_k, we have the past measurement data from the initial time t_0. Using the whole measurement data makes the size of the estimation problem increase with time. To address this, the MHE formulation with finite horizon N is introduced with the arrival cost, which approximates the estimation errors from initial time to t_{k-N}, along with the *prior* estimate of x_{k-N} and its covariance. There are several approaches to calculate the arrival costs such as using EKF, unscented KF, or using the information from the NLP solution (Negrete, 2011; Zavala et al., 2007). The last approach is used in this study.

The ideal MHE formulation for estimating the states from the measurement data is as follows
$\mathcal{M}_N(\eta_k)$:

$$\begin{aligned} \min_{z_l, w_l} \; & \psi_N(\eta_k) = \Gamma(z_{-N}) + \sum_{l=-N}^{-1} L_{N+l}(z_l, w_l) + L_N(z_0) \\ \text{s.t.} \; & z_{l+1} = f(z_l, w_l), \; l = -N, ..., -1 \\ & z_l \in \mathbb{Z}, w_l \in \mathbb{W}. \end{aligned} \tag{2}$$

Here, the arrival cost is:

$$\Gamma(z_{-N}) = \frac{1}{2}(z_{-N} - \bar{z}_{-N})^\top \bar{\Pi}_{-N}^{-1}(z_{-N} - \bar{z}_{-N}), \tag{3}$$

the stage costs are:

$$\begin{aligned} L_{N+l} &= \frac{1}{2}(y_{k+l} - h(z_l))^\top R_l^{-1}(y_{k+l} - h(z_l)) \\ &\quad + \frac{1}{2} w_l^\top Q_l^{-1} w_l, l = -N, ..., -1 \\ L_N &= \frac{1}{2}(y_k - h(z_0))^\top R_0^{-1}(y_k - h(z_0)) \end{aligned} \tag{4}$$

and the *prior* estimate of x_{k-N} and its covariance are denoted as $\bar{z}_{-N}$ and $\bar{\Pi}_{-N}$, respectively. η_k is the data set, $[\bar{z}_{-N}, \bar{\Pi}_{-N}, y_{k-N}, y_{k-N+1}, ..., y_k]$, which fully defines $\mathcal{M}_N$. When we consider this data set as parameters p, the Karush–Kuhn–Tucker (KKT) conditions can be represented as implicit functions of p:

$$\phi(s(p)) = 0 \tag{5}$$

where $s(p)^\top = [z_{-N}^\top, w_{-N}^\top, \lambda_{-N+1}^\top, ..., z_{-1}^\top, w_{-1}^\top, \lambda_0^\top, z_0^\top]$.

In this study, we investigate the effects of perturbations of p on the optimal MHE solution. To this end, we adapt the general NLP sensitivity theorem to the MHE problem. Consider η_k in $\mathcal{M}_N(\eta_k)$ as the nominal parameter p_0, and assume the following.

Assumption 1. $f(\cdot)$, $\Gamma(\cdot)$, and $L(\cdot)$ are twice continuously differentiable in a neighborhood of the nominal solution $s^*(p_0)$.

Assumption 2. $s^*(p_0)$ satisfies the linear independence constraint qualifications (LICQ), second order sufficient conditions (SOSC), and strict complementary slackness (SC).

Then, by the NLP sensitivity theorem (Fiacco, 1976, 1983), there exists an isolated, continuous, and differentiable solution vector $s^*(p)$ for p in a neighborhood of p_0. In addition, $\frac{\partial s}{\partial p}|_{p_0}$ is bounded and unique. From this, we can apply the implicit function theorem to (5) and it yields:

$$K^*(p_0)\frac{\partial s^*}{\partial p}|_{p_0} = -\frac{\partial \phi(s(p))}{\partial p}|_{p_0}. \tag{6}$$

where $K^*(p_0)$ denotes the KKT matrix of $\mathcal{M}_N(p_0)$. For p in a neighborhood of p_0, the first-order Taylor expansion of $s^*(p)$ is

$$s^*(p) \approx s^*(p_0) + \frac{\partial s^*}{\partial p}|_{p_0}(p - p_0) = \tilde{s}(p). \tag{7}$$

The KKT matrix is nonsingular by Assumption 1 (Nocedal and Wright, 2006) and using (6), we have

$$\begin{aligned} \tilde{s}(p) &= s^*(p_0) - K^*(p_0)^{-1}\frac{\partial \phi(s(p))}{\partial p}|_{p_0}(p - p_0) \\ &= s^*(p_0) + \frac{\partial s}{\partial p}|_{p_0}(p - p_0). \end{aligned} \tag{8}$$

where, $\frac{\partial s}{\partial p} = -K^*(p_0)^{-1}\frac{\partial \phi(s(p))}{\partial p}$ is the sensitivity matrix. Thus, we can approximate the optimal solution of $\mathcal{M}_N(p)$ as $\tilde{s}(p)$, using the optimal solution of $\mathcal{M}_N(p_0)$ and its sensitivity matrix. Moreover, the errors between the approximate solution and the optimal solution are bounded by

$$\|\tilde{s}(p) - s^*(p)\| \leq L_s \|p - p_0\|^2 \tag{9}$$

with a constant $L_s > 0$. Thus $s^*(p) - \tilde{s}(p) = \mathcal{O}(\|p - p_0\|^2)$.

3. FAST AMS-MHE ALGORITHM

We propose an advanced-multi-step MHE algorithm to avoid the online computational load when more than one sampling time is required to solve MHE problem. The nominal MHE problem with the predicted measurements is solved offline beforehand. The optimal solution of MHE with the real measurements is approximated using NLP sensitivity online. The detailed algorithm is as follows:

Offline during t_k - t_{k+N_s}:

(1) Predict the future measurements $\hat{y}_{k+1}, ..., \hat{y}_{k+2N_s-1}$ from

$$\begin{aligned} \bar{x}(k+i) &= f(\bar{x}(k+i-1), \bar{w}(k+i-1)) \\ \hat{y}_{k+i} &= h(\bar{x}_{k+i}) \\ i &= 1, \ldots, 2N_s - 1 \\ x(k) &= \bar{x}(k) \end{aligned} \tag{10}$$

The first $N_s - 1$ disturbances can be approximated by the previous MHE problem solved at t_{k-N_s}, while other disturbances are assumed to be zero. That is, $\bar{w}(k+i-1) = 0, \; i = N_s, \ldots, 2N_s - 1$. $\bar{x}(k)$ denotes the estimated values of states at t_k.

(2) Define the extended data $\bar{\eta}_k = (\bar{z}_{-N}(k), \bar{\Pi}^{-1}_{-N}(k), y(k-N)..., y(k), \hat{y}(k+1), ..., \hat{y}(k+2N_s-1))$ and solve the *extended* problem $\mathcal{M}_{N+2N_s-1}(p_0)$ with $p_0 = \bar{\eta}_k$.

$$
\begin{aligned}
\mathcal{M}_{N+2N_s-1}(p_0):& \\
&\min_{z_l, w_l} \psi_{N+2N_s-1}(p_0) \\
&\text{s.t. } z_{l+1} = f(z_l, w_l),\ l = -N, ..., 2N_s-2 \\
&\qquad z_l \in \mathbb{Z}, w_l \in \mathbb{W}.
\end{aligned} \tag{11}
$$

where,

$$
\begin{aligned}
&\psi_{N+2N_s-1}(p_0) = \Gamma(z_{-N}) + \sum_{l=-N}^{0} L_{N+l}(z_l, w_l) \\
&+ \sum_{l=1}^{2N_s-2} \hat{L}_{N+l}(z_l, w_l) + \hat{L}_{N+2N_s-1}(z_{2N_s-1}) \\
&\Gamma(z_{-N}) = \frac{1}{2}(z_{-N} - \bar{z}_{-N+N_s|k-N_s})^\top \bar{\Pi}^{-1}_{-N+N_s|k-N_s} \\
&\qquad \times (z_{-N} - \bar{z}_{-N+N_s|k-N_s}) \\
&L_{N+l} = \frac{1}{2}(y_{k+l} - h(z_l))^\top R_l^{-1}(y_{k+l} - h(z_l)) \\
&\qquad + \frac{1}{2} w_l^\top Q_l^{-1} w_l,\ l = -N, ..., 0 \\
&\hat{L}_{N+l} = \frac{1}{2}(\hat{y}_{k+l} - h(z_l))^\top R_l^{-1}(\hat{y}_{k+l} - h(z_l)) \\
&\qquad + \frac{1}{2} w_l^\top Q_l^{-1} w_l,\ l = 1, ..., 2N_s-2 \\
&\hat{L}_{N+2N_s-1} = \frac{1}{2}(\hat{y}_{k+2N_s-1} - h(z_{N+2N_s-1}))^\top R^{-1}_{2N_s-1} \\
&\qquad \times (\hat{y}_{k+2N_s-1} - h(z_{N+2N_s-1})).
\end{aligned} \tag{12}
$$

(3) Before t_{k+N_s}, obtain the solution $s^*(\bar{\eta}_k)$ and then calculate and compute the sensitivity matrix $\left.\frac{\partial s}{\partial p}\right|_{\bar{\eta}_k}$.

On-line, at $t_{k+i},\ i = N_s, ..., 2N_s-1$**:**
Note that at this point, the measurements $y(k+j), j = 1, \ldots, N_s-1$ are already obtained while $\mathcal{M}_{N+2N_s-1}(\bar{\eta}_k)$ is solved.

(1) Obtain the measurement $y(k+i)$ and update the *current* problem data $\eta(k+i) = (\bar{z}_{-N}(k), \bar{\Pi}^{-1}_{-N}(k), y(k-N)..., y(k+i), \bar{y}(k+i+1), ...\bar{y}(k+2N_s-1))$.
(2) Using *all* $N+1+i$ measurements in $\eta(k+i)$, compute an instantaneous approximate solution $\tilde{s}(\eta(k+i))$ using (8).
(3) Extract the updated solution $\tilde{z}_i$ corresponding to the nominal optimal solution z_i^*; these are the estimated states, $\bar{x}(k+i) = \tilde{z}_i$.

At t_{k+N_s}**:**

(1) Set k to $k+N_s$, and prepare the next iteration. We set $\bar{w}(k+i-1) = \tilde{w}_i,\ i = N_s, ..., 2N_s-1$,which is extracted from the updated solution $\tilde{s}(\eta(k))$. $\bar{z}_{-N}(k)$ is set as $\tilde{z}_{-N+N_s}$. The arrival cost $\bar{\Pi}^{-1}_{-N}(k)$ is approximated using the inverse of reduced Hessian.

Note that the solution of the MHE problem $\mathcal{M}_{N+2N_s-1}(\bar{\eta}_k)$ initiated at t_k is used to estimate states only between t_{k+N_s} and t_{k+2N_s-1}. The states within t_k and t_{k+N_s-1} are estimated using the previous MHE solution, initiated at t_{k-N_s}.

4. SIMULATION RESULTS

The proposed amsMHE is applied to a CSTR, where the exothermic reaction $A \rightarrow B$ occurs. The dimensionless dynamic models in Hicks and Ray (1971); Yang and Biegler (2013) are used with the state disturbances w_1 and w_2.

$$
\begin{aligned}
\frac{dx_1}{dt} &= \frac{1}{u_2}(1 - x_1) - k' exp(-E'/x_2) x_1^3 + w_{x_1} \\
\frac{dx_2}{dt} &= \frac{1}{u_2}(x_f - x_2) + k' exp(-E'/x_2) x_1^3 \\
&\quad - A_h u_1 (x_2 - x_c) + w_{x_2} \\
y &= x_2
\end{aligned} \tag{13}
$$

x_1 and x_2 are the dimensionless A concentration and temperature, respectively. x_2 is measured. x_f and x_c are the dimensionless feed concentration and temperature, respectively. k' and E' are the dimensionless rate constant and ratio of the activation energy to the gas constant, respectively. A_h is the dimensionless heat transfer area. The first manipulated variable u_1 is the reactor jacket heat transfer coefficient which increases monotonically as the coolant flow rate increases. The second manipulated variable u_2 is V/F where F is the feed flow rate and V is the reactor volume. In addition, model parameters are $x_f = 0.395$, $x_c = 0.382$, $k' = 17328$, $E' = 5$ and $A_h = 1.95 \times 10^{-4}$. The moving horizon length N is set as 20 steps with the sampling time of 1 s. The constraints for the states are $[0, 1]$ because both are dimensionless values.

The continuous model is discretized by Lagrange-Radau collocation method, and it is solved using Pyomo (Hart et al., 2011) and IPOPT 3.12 (Wächter and Biegler, 2006) with the linear solver MA57 (Duff, 2004). The KKT matrix and the inverse of the reduced Hessian are obtained using *k_aug* (Thierry and Biegler, 2019).

4.1 With state disturbances and measurement noise

We simulate four different cases with the state disturbance w and measurement noise v:

- Case 1: $w_{x_1}$and $w_{x_2} \sim \mathcal{N}(0, 0.01^2), v \sim \mathcal{N}(0, 0.01^2)$
- Case 2: $w_{x_1}$and $w_{x_2} \sim \mathcal{N}(0, 0.02^2), v \sim \mathcal{N}(0, 0.01^2)$
- Case 3: $w_{x_1}$and $w_{x_2} \sim \mathcal{N}(0, 0.01^2), v \sim \mathcal{N}(0, 0.02^2)$
- Case 4: $w_{x_1}$and $w_{x_2} \sim \mathcal{N}(0, 0.02^2), v \sim \mathcal{N}(0, 0.02^2)$

We use the fixed control inputs obtained by the ideal NMPC and the ideal MHE for the asMHE ($N_s = 1$) and amsMHE ($Ns = 3$) simulation. The R_l and Q_l are set as the standard deviation of each noise.

The results are shown in Fig. 1 and Table 1. The sum-of-squared errors for asMHE are smaller than for ideal MHE. Although the sum-of-squared errors in x_2 for cases 3 and 4 are smaller for amsMHE, the total sum-of-squared errors are always greater than those of ideal MHE.

To investigate why asMHE can yield better estimations than ideal MHE, we analyze the KKT conditions of ideal MHE and the approximate update of asMHE. This occurs because the MHE objective function contains the state disturbance terms. In section 4.2, we show the results of the case without state disturbances, where the objective function does not include them and the ideal MHE estimates the true values better than asMHE.

Consider the KKT conditions of the ideal MHE at t_{k+1} where the data set $\eta_{k+1} = [\bar{z}_{-N+1}, \bar{\Pi}_{-N+1}, y_{k-N+1}, ..., y_{k+1}]$ is used with N horizon, and the inequality constraints are assumed inactive. The Lagrangian of $\mathcal{M}_N(\eta_{k+1})$ is

$$\mathcal{L} = \psi_N(\eta_{k+1}) + \sum_{l=-N+1}^{0} \lambda_{l+1}^\top (z_{l+1} - f(z_l, w_l)). \tag{14}$$

and the KKT conditions for z_1 and λ_1 are,

$$\nabla_{z_1}\mathcal{L} = -\nabla_{z_1} h_1^\top R_1^{-1}(y_{k+1} - h(z_1)) + \lambda_1 = 0 \tag{15a}$$
$$\nabla_{\lambda_1}\mathcal{L} = z_1 - f(z_0, w_0) = 0 \tag{15b}$$
$$\nabla_{w_0}\mathcal{L} = Q_0^{-1}w_0 - \nabla_{w_0} f_0^\top \lambda_1 = 0. \tag{15c}$$

Here, $h_l = h(z_l)$ and $f_l = f(z_l, w_l)$. The solution of ideal MHE z_1^*, w_0^*, and λ_1^* satisfies (15). In addition, because $h(z_1) = [0, 1][x_1(t_{k+1}), x_2(t_{k+1})]^\top$ for this example, (15a) is simplified as

$$\lambda_1^* = \begin{bmatrix} 0 \\ R_1^{-1}(y_{k+1} - x_2^*(t_{k+1})) \end{bmatrix} \tag{16}$$

With $\bar{\eta}_k = [\bar{z}_{-N}, \bar{\Pi}_{-N}, y_{k-N+1}, ..., y_k, \hat{y}_{k+1}]$ and an extended horizon $N+1$, the Lagrangian of asMHE becomes:

$$\mathcal{L}_{as} = \psi_{N+1}(\bar{\eta}_k) + \sum_{l=-N}^{0} \lambda_{l+1}^\top (z_{l+1} - f(z_l, w_l)). \tag{17}$$

The solution $z_{1,as}^*$, $w_{0,as}^*$, and $\lambda_{1,as}^*$ of $\mathcal{M}_{N+1}(\bar{\eta}_k)$ satisfies the following parts of the KKT conditions:

$$\nabla_{z_1}\mathcal{L}_{as} = -\nabla_{z_1} h_1^\top R_1^{-1}(\hat{y}_{k+1} - h(z_1)) + \lambda_1 = 0 \tag{18a}$$
$$\nabla_{\lambda_1}\mathcal{L} = z_1 - f(z_0, w_0) = 0 \tag{18b}$$
$$\nabla_{w_0}\mathcal{L} = Q_0^{-1}w_0 - \nabla_{w_0} f_0^\top \lambda_1 = 0. \tag{18c}$$

When the measurement y_{k+1} is obtained, the solution is updated to satisfy

$$\nabla_{z_1 z_1}\mathcal{L}_{as}\Delta z_1 + \Delta\lambda_1 - \nabla_{z_1} h_1^\top R_1^{-1}\Delta p = 0 \tag{19}$$

where the variables $\Delta z_1 = \tilde{z}_1 - z_{1,as}^*$, $\Delta\lambda = \tilde{\lambda}_1 - \lambda_{1,as}^*$, and $\Delta p = y_{k+1} - \hat{y}_{k+1}$. Because $h(z_1) = [0, 1][x_1(t_{k+1}), x_2(t_{k+1})]^\top$ in this example, (18a) becomes

$$-\begin{bmatrix} 0 \\ 1 \end{bmatrix} R_1^{-1}(\hat{y}_{k+1} - x_{2,as}^*(t_{k+1})) + \lambda_{1,as}^* = 0 \tag{20}$$

and (19) becomes

$$\begin{bmatrix} 0 \\ R_1^{-1}(\tilde{x}_2(t_{k+1}) - x_{2,as}^*(t_{k+1})) \end{bmatrix} + \Delta\lambda_1 - \begin{bmatrix} 0 \\ R_1^{-1}\Delta p \end{bmatrix} = 0 \tag{21}$$

From (20) and (21),

$$\begin{bmatrix} 0 \\ R_1^{-1}(\tilde{x}_2(t_{k+1})) \end{bmatrix} + \tilde{\lambda}_1 - \begin{bmatrix} 0 \\ R_1^{-1}y_{k+1} \end{bmatrix} = 0. \tag{22}$$

and

$$\tilde{\lambda}_1 = \begin{bmatrix} 0 \\ R_1^{-1}(y_{k+1} - \tilde{x}_2(t_{k+1})) \end{bmatrix}. \tag{23}$$

Considering (15c) and (18c) and comparing (16) and (23), we can note that the estimates of $x_2(t_{k+1})$ depend on the estimates of state disturbances w_0 in addition to z_0. This may make the asMHE estimate of $x_2(t_{k+1})$ better from the perspective of the errors.

On the other hand, as seen in Table 1 for amsMHE with $N_s = 3$, solving the optimization problem less frequently, and with less accurate predicted measurements, leads to a deterioration of estimation performance. In addition, the amsMHE approximates an ideal MHE with real and predicted measurements depending on the time steps. This can also affect estimation performance.

Table 1. The squared errors between the estimates and true values of amsMHE with $N_s = 3$, asMHE, and ideal MHE when considering state disturbances and measurement noise.

		amsMHE	asMHE	Ideal MHE
Case 1	$\sum_0^{150}(x_1 - \hat{x}_1)^2$	0.0473	0.0081	0.0102
	$\sum_0^{150}(x_2 - \hat{x}_2)^2$	0.0601	0.0113	0.0541
	Total	0.1074	0.0194	0.0644
Case 2	$\sum_0^{150}(x_1 - \hat{x}_1)^2$	0.1965	0.0293	0.0319
	$\sum_0^{150}(x_2 - \hat{x}_2)^2$	0.1988	0.0147	0.1653
	Total	0.3953	0.044	0.1972
Case 3	$\sum_0^{150}(x_1 - \hat{x}_1)^2$	0.0389	0.0121	0.0168
	$\sum_0^{150}(x_2 - \hat{x}_2)^2$	0.0579	0.0292	0.0728
	Total	0.0968	0.0413	0.0895
Case 4	$\sum_0^{150}(x_1 - \hat{x}_1)$	0.1684	0.0353	0.0423
	$\sum_0^{150}(x_2 - \hat{x}_2)$	0.1758	0.0447	0.1914
	Total	0.3442	0.08	0.2338

4.2 Without state disturbances and with measurement noise

To further compare the effects of the approximate approach, we simulate the MHE without state disturbances and only with measurement noise. The results obtained with $v \sim \mathcal{N}(0, 0.05^2)$ are shown in Table 2 and Fig. 2. As expected, the performances of amsMHE deteriorate as N_s increases because the NLP is solved with a greater number of predicted measurements and then the solution is updated approximately when the real measurements are obtained.

In Table 3, we show the average and maximum of online and offline CPU time to solve MHEs. The online CPU time of asMHE is negligible and the average and maximum online CPU times of amsMHE with $N_s = 3$ are smaller than those of ideal MHE. Because the example has small computational load, the time reductions for asMHE and amsMHE are also small. A detailed large-scale case study with an ASU model (Kim et al. (2020a), submitted) shows a reduction in average on-line computation from over 100 CPU s to about 0.001 CPU s.

Table 2. The squared errors between the estimates and true values of amsMHE with $N_s = 3$, asMHE, and ideal MHE for the simulations only with measurement noise.

	amsMHE	asMHE	Ideal MHE
$\sum_0^{150}(x_1 - \hat{x}_1)^2$	0.009	8.11E-06	2.65E-07
$\sum_0^{150}(x_2 - \hat{x}_2)^2$	0.0263	0.0014	9.50E-07
Total	0.0353	0.0014	1.22E-06

5. CONCLUSION

This paper proposes an advanced-multi-step MHE that the optimization problem with predicted outputs is solved in advance, and the optimal estimates with the real measurements are approximated by NLP sensitivity update online.

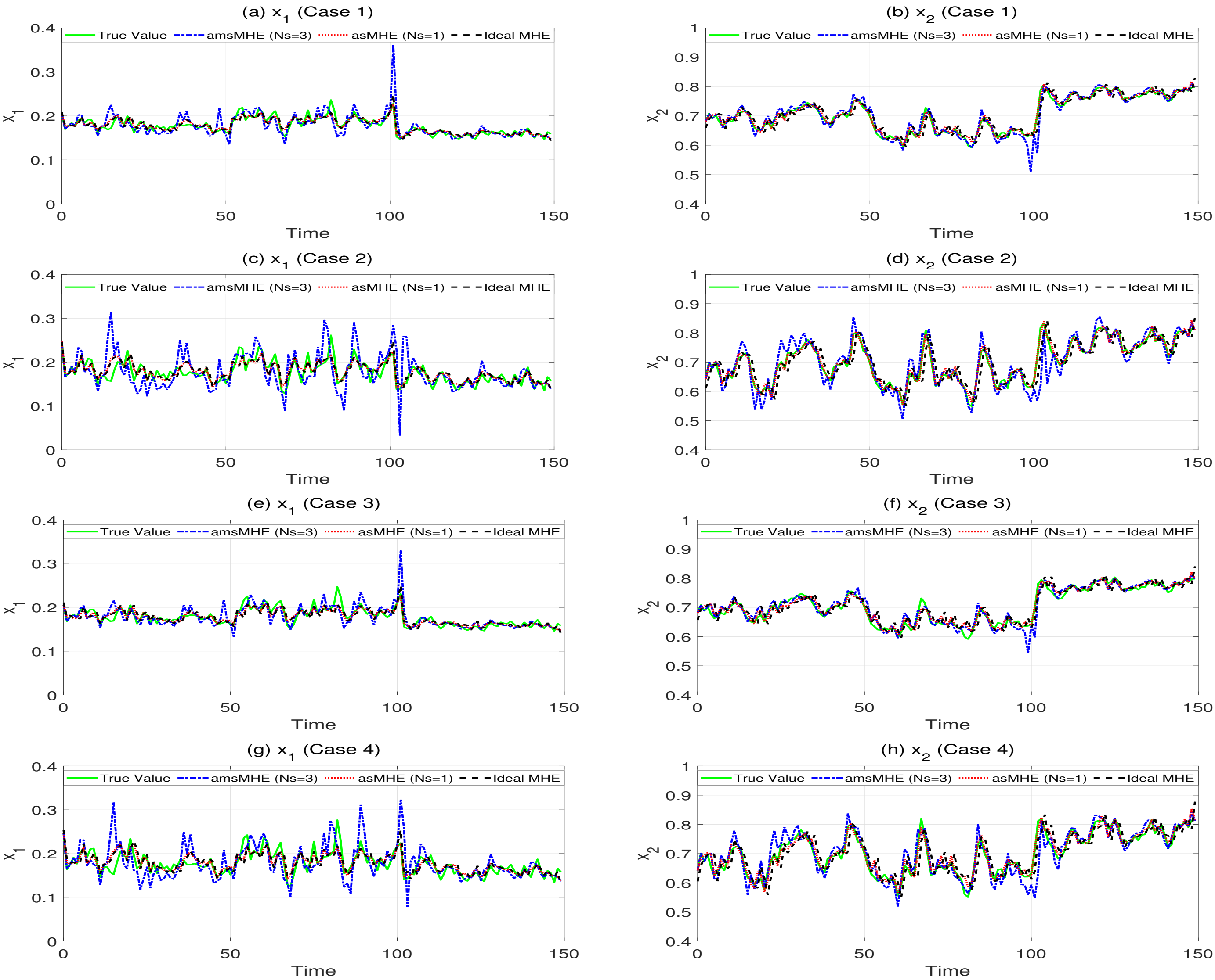

Fig. 1. State estimation with state disturbances and measurement noise.

Table 3. Average and maximum CPU times for solving ideal MHE and amsMHEs.

		amsMHE	asMHE	Ideal MHE
Offline CPU time [s]	Average	0.0042	0.0070	-
	Max	0.0316	0.0456	-
Online CPU time [s]	Average	0.0005	0.0000	0.0055
	Max	0.0156	0.0000	0.0160

We investigate the performances of amsMHE, asMHE and ideal MHE with state disturbances and measurement noise. The estimation performance of asMHE is sometimes better than ideal MHE because general MHE problems minimize the sum of estimation errors and the norm of state disturbances. On the other hand, given the stochastic nature of these simulations, it is expected that performance comparisons among these approaches may be case dependent. Also, when we simulate the cases only with measurement noise, we observe that the performance deteriorates as N_s increases. The computational load of NLP can be handled by ams strategy, solving the optimization problem in background; thus, future work will consider large-scale problems to show the computational advantages of the amsMHE method. Also, parametric uncertainties will be considered, and the amsMHE method will be coupled with advanced-multi-step nonlinear model predictive control (amsNMPC), thus leading to a fast output-based controller.

ACKNOWLEDGEMENTS

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education [2019R1A6A3A12031355]. Kuan-Han Lin, David Thierry and Lorenz Biegler also acknowledge support from the Center for Advanced Process Decision-Making (CAPD) at Carnegie Mellon University.

REFERENCES

Bryson, A. and Ho, Y.C. (1975). *Applied Optimal Control: Optimization, Estimation, and Control.* CRC Press, New York.

Chen, W., Ballance, D.J., and O'Reilly, J. (2000). Model predictive control of nonlinear systems: computational

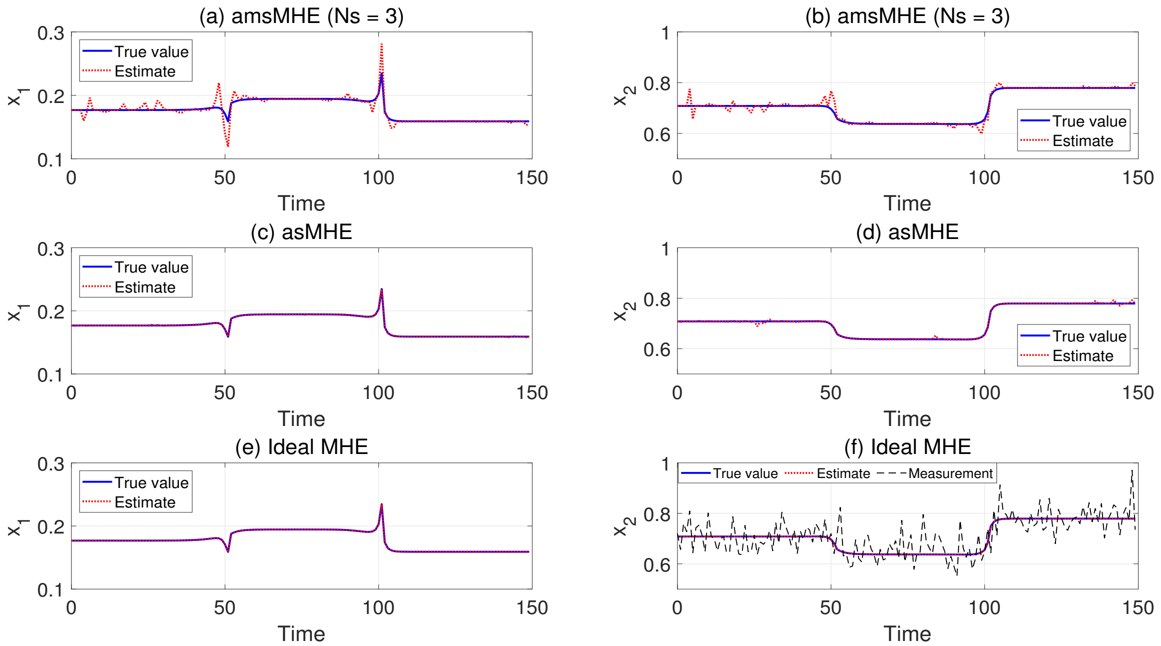

Fig. 2. State estimation without state disturbances and only with measurement noise.

burden and stability. *IEE Proceedings - Control Theory and Applications*, 147(4), 387–394.

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

Duff, I.S. (2004). MA57 – a code for the solution of sparse symmetric definite and indefinite systems. *ACM Transactions on Mathematical Software*, 30(2), 118–144.

Fiacco, A.V. (1976). Sensitivity analysis for nonlinear programming using penalty methods. *Mathematical Programming*, 10(1), 287–311.

Fiacco, A.V. (1983). *Introduction to sensitivity and stability analysis in nonlinear programming*. Academic Press.

Findeisen, R. and Allgöwer, F. (2004). Computational delay in nonlinear model predictive control. *IFAC Proceedings Volumes*, 37(1), 427–432.

Hart, W.E., Watson, J.P., and Woodruff, D.L. (2011). Pyomo: modeling and solving mathematical programs in python. *Mathematical Programming Computation*, 3, 219–260.

Haseltine, E.L. and Rawlings, J.B. (2005). Critical evaluation of extended kalman filtering and moving-horizon estimation. *Industrial & Engineering Chemistry Research*, 44(8), 2451–2460.

Hicks, G.A. and Ray, W.H. (1971). Approximation methods for optimal control synthesis. *The Canadian Journal of Chemical Engineering*, 49(4), 522–528.

Jazwinski, A. (1970). *Stochastic Processes and Filtering Theory*. Academic Press, New York.

Kim, Y., Lin, K.H., Thierry, D.M., and Biegler, L.T. (2020a). Advanced-multi-step MHE for large-scale nonlinear systems. *Submitted for publication*.

Kim, Y., Thierry, D.M., and Biegler, L.T. (2020b). Serial advanced-multi-step nonlinear model predictive control using an extended sensitivity method. *Journal of Process Control*, 96, 82 – 93.

Kraus, T., Kuhl, P., Wirsching, L., Bock, H.G., and Diehl, M. (2006). Moving Horizon State Estimation for Tennessee Eastman Benchmark Process. In *2006 IEEE Intl. Conf. on Multisensor Fusion and Integration for Intelligent Systems*, 377–382.

Negrete, R.L. (2011). *Nonlinear Programming Sensitivity Based Methods for Constrained State Estimation*. Ph.D. thesis, Carnegie Mellon University.

Nocedal, J. and Wright, S. (2006). *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York.

Thierry, D.M. and Biegler, L.T. (2019). Dynamic real-time optimization for a CO2 capture process. *AIChE Journal*, 65(7), 1–11.

Wächter, A. and Biegler, L.T. (2006). Implementation of an interior-point filter algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1), 25–57.

Wynn, A., Vukov, M., and Diehl, M. (2014). Convergence guarantees for moving horizon estimation based on the real-time iteration scheme. *IEEE Transactions on Automatic Control*, 59(8), 2215–2221.

Yang, X. and Biegler, L.T. (2013). Advanced-multi-step nonlinear model predictive control. *Journal of Process Control*, 23(8), 1116–1128.

Zavala, V.M., Laird, C.D., and Biegler, L.T. (2007). A fast computational framework for large-scale moving horizon estimation. *IFAC Proceedings Volumes*, 40(5), 19 – 28. 8th IFAC Symposium on Dynamics and Control of Process Systems.

Zavala, V.M., Laird, C.D., and Biegler, L.T. (2008). A fast moving horizon estimation algorithm based on nonlinear programming sensitivity. *Journal of Process Control*, 18(9), 876 – 884.