Robust Moving Horizon State Estimation for Nonlinear Systems

Jinfeng Liu

Abstract—In this work, a robust moving horizon estimation scheme augmented by an auxiliary nonlinear observer is proposed for nonlinear systems with bounded model uncertainties. Specifically, an auxiliary deterministic nonlinear observer that asymptotically tracks nominal system state is taken advantage of to calculate a confidence region that contains the actual system state taking into account the effects of bounded model uncertainties at every sampling time. This region is then used to design a constraint on the state estimates in the proposed moving horizon estimation. The proposed design brings together deterministic and optimization-based observer design techniques. First, the proposed moving horizon estimation scheme is proved to give bounded estimation errors in the case of bounded model uncertainties. Second, the proposed approach provides another option to compromise the effects of errors in arrival cost approximations and can be used together with different arrival cost approximation techniques to further improve the state estimate.

I. INTRODUCTION

State estimation plays an important role in feedback control, system monitoring, fault detection as well as system optimization because it is in general difficult to measure all the system state variables. Observer designs that explicitly account for nonlinear systems can track back to the earlier work by Thau [10] in which a sufficient condition for nonlinear observer asymptotical convergence was proposed. In [4], the first systematic approach for the design of nonlinear observers was proposed in which a nonlinear state transformation is used to linearize the original nonlinear system to a form that linear methods can be applied. Even though an effective design for general nonlinear systems is still not available, significant progress has been made in the design of observers for different specific classes of nonlinear systems (e.g., [1], [11], [2]) with many successful applications to nonlinear chemical processes (e.g., [3], [8]). Important features of these nonlinear observer designs are that the observers can be expressed analytically and possess stability and convergence properties that can be proved rigorously. These observers can be classified as deterministic since, in the design of these observers, model uncertainties and measurement noise are not taken into account, which may lead to deteriorated performance or even lose of stability of the observer. In addition, only the current measurement is used in these observers which makes the observers sensitive to occasional spikes in process disturbances and measurement noise.

Moving horizon estimation (MHE) based on batch least squares techniques has become popular in recent years because of its ability to handle explicitly nonlinear systems and constraints on decision variables (e.g., [5], [6], [7]). In MHE, the state estimate is determined by solving online an optimization problem which minimizes the sum of squared errors. At a sampling time, when a new measurement is available, the oldest measurement in the estimation window is discarded, and the finite horizon optimization problem is solved again to get the new estimate of the state [5]. The ability of MHE to handle constraints on process disturbances, measurement noise and states was shown to lead to improved performance [6]. In order to account for the effect of historical data outside the estimation window, an arrival cost which summarizes the information of those data is included in the cost function of the MHE optimization problem. The arrival cost plays a critical role in the performance and stability of MHE which makes the approximation of the arrival cost an important issue. On the other hand, the incorporation of constraints in the optimization problem and the use of finite estimation horizon pose difficulties in the associated theory development of MHE especially when bounded uncertainties are involved [7].

Motivated by the above observations and inspired by techniques developed in nonlinear model predictive control with Lyapunov-based stability constraints, an MHE scheme augmented by an auxiliary nonlinear observer is proposed for nonlinear systems with bounded model uncertainties in this work. Specifically, a deterministic nonlinear observer that asymptotically tracks the nominal system state is used to calculate a confidence region that contains the actual system state taking into account the effects of bounded model uncertainties at every sampling time. This confidence region is then used to design a constraint on state estimates in the MHE scheme. The proposed MHE scheme brings together deterministic and optimization-based observer design techniques. First of all, the proposed MHE scheme is proved to give bounded estimation errors. Second, it is demonstrated through the application to a gas-phase reactor that the proposed MHE scheme is less sensitive to the accuracy of the approximated arrival cost compared with the classical MHE schemes. This proposed approach gives us another option to compensate for the effects of errors in the arrival cost approximation, and it can be used together with different arrival cost approximation techniques to further improve state estimates.

II. PRELIMINARIES

A. Notation

The operator $| \cdot |$ is used to denote Euclidean norm of a scalar or a vector while $| \cdot |^2_Q$ indicates the square of the
weighted Euclidean norm of a vector, defined by \( |x|^2_Q = x^T Q x \). A function \( f(x) \) is said to be locally Lipschitz with respect to its argument if there exists a constant \( L_f^x \) such that \( |f(x') - f(x'')| \leq L_f^x |x' - x''| \) for all \( x' \) and \( x'' \) in a given region of \( x \) and \( L_f^x \) is the associated Lipschitz constant. A continuous function \( \gamma : [0, a) \to [0, \infty) \) is said to belong to class \( \mathcal{K} \) if it is strictly increasing and satisfies \( \gamma(0) = 0 \).

A function \( \beta(r, s) \) is said to be a class \( \mathcal{KL} \) function if, for each fixed \( s \), \( \beta(r, s) \) belongs to class \( \mathcal{K} \) function with respect to \( r \) and, for each fixed \( r \), \( \beta(r, s) \) is decreasing with respect to \( s \) and \( \beta(r, s) \to 0 \) as \( s \to 0 \). Symbols with a subscript 'n' are associated with nominal system/observer states and noise free measurements.

**B. System description**

We consider nonlinear systems with the following state-space representation:

\[
\begin{align*}
\dot{x}(t) &= f(x(t), w(t)) \\
x(0) &= x_0 \\
y(t) &= h(x) + v(t)
\end{align*}
\]

where \( x \in \mathbb{R}^n \) is the state vector with \( x_0 \) the initial state at time \( t = 0 \), \( w \in \mathbb{R}^p \) is a state disturbance vector, \( y \in \mathbb{R}^m \) is the measured output and \( v \in \mathbb{R}^q \) is a measurement noise vector. In order to simplify the discussion but without loss of generality, inputs of the system are not considered.

We assume that the state \( x \) satisfies the constraint:

\[ x \in \mathcal{X} \]

where \( \mathcal{X} \) is a convex compact set and the disturbances are bounded such as \( w \in \mathcal{W} \) and \( v \in \mathcal{V} \) where

\[
\begin{align*}
\mathcal{W} &:= \{ w \in \mathbb{R}^p : |w| \leq \theta, \ \theta > 0 \} \\
\mathcal{V} &:= \{ v \in \mathbb{R}^q : |v| \leq \theta_v, \ \theta_v > 0 \}
\end{align*}
\]

with \( \theta \) and \( \theta_v \) known positive real numbers. We also assume that \( f \) is locally Lipschitz of its arguments in \( \mathcal{X} \times \mathcal{W} \) and \( h \) is locally Lipschitz of its argument in \( \mathcal{X} \) with \( f(0,0) = 0, h(0) = 0 \). We further assume that the output of system (1), \( y \), is sampled at synchronous time instants \( \{t_k \geq 0\} \) such that \( t_k = t_0 + k\Delta \) with \( t_0 = 0 \) the initial time, \( \Delta \) a fixed time interval and \( k \) positive integers.

**C. Deterministic nonlinear observer**

We first define the nominal system of system (1) with noise free measurements as follows:

\[
\begin{align*}
\dot{x}_n(t) &= f(x_n(t), 0) \\
x_n(0) &= x_{n0} \\
y_n(t) &= h(x_n)
\end{align*}
\]

where \( x_n \in \mathcal{X} \subset \mathbb{R}^n \) is the state of system (2) with \( x_{n0} \) the initial state and \( y_n \in \mathbb{R}^m \) is the noise free output of system (2). We assume that there exists a deterministic nonlinear observer for the nominal system (2) which takes the following form:

\[
\begin{align*}
\dot{z}_n(t) &= F(z_n(t), y_n(t)) \\
z_n(0) &= z_{n0}
\end{align*}
\]

such that \( z_n \) asymptotically approaches \( x_n \) for all the states \( z_n, x_n \in \mathcal{X} \). This assumption implies that there exists a \( \mathcal{KL} \) function \( \beta \) such that:

\[ |z_n(t) - x_n(t)| \leq \beta(|z_{n0} - x_{n0}|, t). \]  

We also assume that \( F \) is locally Lipschitz of its arguments. Note that the convergence property of observer (3) is obtained based on nominal system (2) with continuous output feedback.

Based on the Lipschitz property assumed for the vector field \( f \), there exists a positive constant \( M \) such that:

\[ |f(x_n, 0)| \leq M \]  

for all \( x_n \in \mathcal{X} \). This constant together with the Lipschitz constants associated with \( f, F, h \) will be used to characterize the properties of the nonlinear observer.

**III. PROPERTIES OF THE NONLINEAR OBSERVER**

In this section, we study the robustness properties of the deterministic observer (3) when it is applied to system (1) with sampled output measurements and the presence of process disturbance and measurement noise. These properties will be taken advantage of in the proposed MHE design which will be discussed in Section IV. In particular, we investigate that how much the estimated state provided by observer (3) will deviate from the actual system state in one sampling time if only sampled (not continuous) output measurements are available and process disturbance and measurement noise are present. In this case, the observer operates in an open-loop manner between two measurement samples.

In order to distinguish from the nominal closed-loop case, we use \( z \) to denote the state of observer (3) when it is applied to system (1) with sampled output measurements. Specifically, from \( t_k \) to \( t_{k+1} \), the dynamics of \( z \) is described as follows:

\[
\begin{align*}
\dot{z}(t) &= F(z(t), y(t)) \\
z(t_k) &= \hat{x}(t_k)
\end{align*}
\]

where \( y(t_k) \) is the actual sampled measurement at \( t_k \) and \( \hat{x}(t_k) \) is an estimate of the actual system state \( x(t_k) \). The following proposition provides an upper bound on the deviation of \( z \) from \( x \).

**Proposition 1:** Consider observer (6) applied to system (1) with sampled output measurement \( y(t_k) \) starting from the initial condition \( \hat{x}(t_k) \), the deviation of the observer state \( z \) from the actual system state \( x \) is bounded for \( t \in [t_k, t_{k+1}] \) as follows:

\[ |z(t) - x(t)| \leq \beta(|e(t_k)|, t-t_k) + \gamma_x(t-t_k) + \gamma_z(t-t_k) \]

for \( z, x \in \mathcal{X} \) where \( e(t) = \hat{x}(t) - x(t) \) is the difference between the estimated state and the actual system state and

\[
\begin{align*}
\gamma_x(\tau) &= \frac{L^x}{L_f} \left( e^{L_f \tau} - 1 \right), \\
\gamma_z(\tau) &= \frac{L^z}{L_f} \left( L_h M \Delta + \theta_v \right) \left( e^{L_f \tau} - 1 \right)
\end{align*}
\]
with \( L^x_f \), \( L^z_f \), \( L^y_f \), and \( L^y_h \) are Lipschitz constants associated with \( f \), \( F \) and \( h \), respectively, for \( x, z \in \mathbb{R} \).

**Proof:** Consider the evolutions of the following four systems in the time interval \( t \in [t_k, t_{k+1}] \): system (1) starting from \( x(t_k) \) (i.e., the actual system), system (2) starting from \( x(t_k) \) (i.e., the nominal system), observer (3) starting from \( \hat{x}(t_k) \) (i.e., the nominal observer), and observer (4) starting from \( \hat{x}(t_k) \) (i.e., the actual observer).

From condition (4) and the fact that \( z_n(t_k) = \hat{z}(t_k) \) and \( x_n(t_k) = x(t_k) \), the following inequality can be obtained for \( t \in [t_k, t_{k+1}] \):

\[
|z_n(t) - x_n(t)| \leq \beta(\hat{z}(t_k) - x(t_k)), t-t_k. \tag{9}
\]

From the Lipschitz property of \( f \), there exists Lipschitz constants \( L^x_f \), \( L^y_f \) such that the following inequality holds:

\[
|\dot{x}_n(t) - \dot{x}(t)| \leq L^x_f |x_n(t) - x(t)| + L^y_f |w(t)|. \tag{10}
\]

Taking into account that \( x_n(t_k) = x(t_k) \) (i.e., \( |x_n(t_k) - x(t_k)| = 0 \)) and \( |w(t)| \leq \theta \), integrate inequality (10) from \( t_k \) to \( t \), the following inequality can be obtained:

\[
|z_n(t) - x(t)| \leq \beta(\hat{z}(t_k) - x(t_k)) + \gamma_x(t-t_k) \tag{11}
\]

with \( \gamma_x(t) \) defined in (8).

From the triangle inequality, it can be written that:

\[
|z_n(t) - x(t)| \leq |z_n(t) - x_n(t)| + |x_n(t) - x(t)|. \tag{12}
\]

From inequalities (9), (11) and (12), it can be obtained that for \( t \in [t_k, t_{k+1}] \):

\[
|z_n(t) - x(t)| \leq \beta(\hat{z}(t_k) - x(t_k)) + \gamma_x(t-t_k) \tag{13}
\]

where \( e(t) := \hat{x}(t) - x(t) \).

Define \( e_z(t) := z(t) - z_n(t) \) and from the dynamics of the observer, it can be written that:

\[
\dot{e}_z(t) = F(z(t), y(t)) - F(z_n(t), y_n(t)) \tag{14}
\]

for \( t \in [t_k, t_{k+1}] \). From the Lipschitz property of \( F \), there exists Lipschitz constants \( L^z_F \) and \( L^y_F \) such that the following inequality holds:

\[
|\dot{e}_z(t)| \leq L^z_F |e_z(t)| + L^y_F |y(t) - y_n(t)| \tag{15}
\]

for \( t \in [t_k, t_{k+1}] \). Noticing that \( y(t_k) = h(x(t_k)) + v(t_k) \) and \( y_n(t) = h(x_n(t_k)) \), it can be written that:

\[
|y(t_k) - y_n(t)| \leq |h(x(t_k)) - h(x_n(t_k))| + |v(t)|. \tag{16}
\]

From the Lipschitz property of \( h \) and inequality (16), there exists a Lipschitz constant \( L_h \) satisfies the following inequality:

\[
|y(t_k) - y_n(t)| \leq L_h |x(t_k) - x_n(t)| + |v(t)|. \tag{17}
\]

Recalling that \( x(t_k) = x_n(t_k) \) and that \( v \) is bounded, from (17), it can be obtained that:

\[
|y(t_k) - y_n(t)| \leq L_h |x_n(t_k) - x_n(t)| + \theta_v. \tag{18}
\]

From (5) and the dynamics of \( x_n \), it can be obtained that:

\[
|x_n(t_k) - x_n(t)| \leq M(t-t_k). \tag{19}
\]

From (18) and (19), it can be derived that:

\[
|y(t_k) - y_n(t)| \leq L_h M(t-t_k) + \theta_v. \tag{20}
\]

From (15) and (20) and taking into account that \( |t-t_k| \leq \Delta \) for \( t \in [t_k, t_{k+1}] \), the following inequality can be obtained:

\[
|\dot{e}_z(t)| \leq L^z_F |e_z(t)| + L^y_F L_h |z(t)| + L^y_F \theta_v \tag{21}
\]

for \( t \in [t_k, t_{k+1}] \). From (21) and take into account that \( e_z(t_k) = z(t_k) - z_n(t_k) = 0 \), the following condition is obtained:

\[
|\dot{e}_z(t)| \leq \gamma_z(t-t_k) \tag{22}
\]

with \( \gamma_z(\cdot) \) defined in (8).

From the triangle inequality, it can be written that:

\[
|z(t) - x(t)| \leq |z(t) - z_n(t)| + |z_n(t) - x(t)|. \tag{23}
\]

Taking into account (13), (22) and (23), inequality (7) is obtained for \( t \in [t_k, t_{k+1}] \). This proves Proposition 1.

Proposition 1 provides an upper bound on the deviation of the open-loop deterministic observer state \( z \) from the actual system state \( x \). From the expression of (7), it can be seen that the value of this upper bound depends on several factors including the accuracy of the initial state estimate (i.e., \( c(t_k) \)), Lipschitz properties of the system and observer dynamics, sampling time of measurements (i.e., \( \Delta \)), magnitudes of disturbances and noise (i.e., \( \theta \) and \( \theta_v \)) as well as open-loop operation time (i.e., \( \Delta - t_k \)). The boundedness property of the deviation will be utilized to design a constraint in the proposed MHE scheme which gives bounded estimation errors.

**IV. PROPOSED MOVING HORIZON ESTIMATION**

In this section, a new moving horizon estimation scheme is proposed for system (1) which includes a novel constraint on the state estimate that takes advantage of the robustness property of the deterministic nonlinear observer. The proposed MHE scheme for time instant \( t_k \) is formulated as follows:

\[
\min_{\hat{x}(t_{k-N}), \ldots, \hat{x}(t_k)} \left\{ \sum_{i=k-N}^{k-1} \left| w(t_i) \right|^2 + \sum_{i=k-N}^{k} \left| v(t_i) \right|^2 + V(t_{k-N}) \right\} \tag{24a}
\]

s.t. \( \hat{x}(t) = f(\hat{x}(t), w(t)) \), \( t \in [t_i, t_{i+1}] \)

\[
v(t_i) = y(t_i) - h(\hat{x}(t_i)) \tag{24b}
\]

\[
w(t_i) \in W, v(t_i) \in V, \hat{x}(t) \in \mathbb{X} \tag{24c}
\]

\[
z(t) = F(z(t), y(t_{k-1})) \tag{24d}
\]

\[
z(t_{k-1}) = \hat{x}(t_{k-1}) \tag{24e}
\]

\[
|\hat{x}(t_k) - z(t_k)| \leq \kappa |y(t_k) - h(z(t_k))| \tag{24f}
\]

where \( N \) is the estimation window size (i.e., horizon), \( Q \) and \( R \) are the covariance matrices of \( w \) and \( v \) respectively, \( V(t_{k-N}) \) denotes the arrival cost which summarizes past information up to \( t_{k-N} \), \( \hat{x} \) is the predicted \( x \) in the above optimization problem, \( \hat{x}(t_k) \) is the optimal estimate of \( x \) at time \( t_{k-1} \), \( y(t_i) \) is the output measurements at \( t_i \), and \( \kappa \) is a positive constant which is a design parameter.
In optimization problem (24), the decision variables are system state estimates \( \hat{x}(t_k-N), \ldots, \hat{x}(t_k) \) which is equivalent to have \( \hat{x}(t_k-N), w(t_k-N), \ldots, w(t_k-1) \) as decision variables. Once optimization problem (24) is solved, an optimal trajectory of the system states, \( \hat{x}^{*}(t_k-N), \ldots, \hat{x}^{*}(t_k) \), is obtained. In state filtering applications (for example, real-time output feedback control), only the estimate of the current system state is used. That is,
\[
\hat{x}(t_k) = \hat{x}^{*}(t_k).
\]
This is also assumed to be the case in this work. Note that in the optimization problem (24), \( w \) and \( v \) are assumed to be piece-wise constant variables with sampling time \( \Delta \) to ensure that (24) is a finite dimensional optimization problem.

Constraint (24a) is the cost function that needs to be minimized. It is in general a difficult task to determine the arrival cost for constrained nonlinear systems. A common approach is to calculate the arrival cost by approximating the constrained nonlinear system as an unconstrained linear time-varying system and computing the arrival cost using extended Kalman filtering scheme. When such approximation is used, the arrival cost \( V(\hat{x}(t_k-N)) \) takes the following form:
\[
V(t_k-N) = |\hat{x}(t_k-N) - x_E(t_k-N)|^2 P(t_k-N)^{-1} \tag{25}
\]
where \( \hat{x}(t_k-N) \) is the state estimate generated in (24), \( x_E(t_k-N) \) is the a priori estimate of \( x(t_k-N) \) predicted based on the process model and the state estimate at \( t_k-N-1 \), the covariance matrix \( P(t_k-N) \) is an estimate of the covariance of \( x_E(t_k-N) \) based on extended Kalman filtering. Poor estimates of the arrival cost may lead to deteriorated performance or even loss of stability of the classical MHE scheme and the use of long horizon increases the computational complexity of MHE.

Constraints (24b) and (24c) are system model governing the evolution of the system states and outputs. Constraint (24d) are the constraints on disturbances, measurement noise and system states. These constraints are also used in the classical MHE scheme.

The novelty and the main different of the proposed MHE scheme from the classical MHE scheme are in constraints (24e)-(24g). The idea of these constraints is to construct a confidence region that contains the actual system state based on the information provided by the deterministic nonlinear observer. In the proposed MHE scheme, the estimate of the current state is only allowed to be optimized within the confidence region. This approach ensures that the proposed MHE gives estimates with bounded errors and is robust with respect to poor arrival cost approximations. Specifically, constraints (24e) and (24f) are the deterministic nonlinear observer with sampled measurement \( y(t_k-1) \) starting from the optimal state estimate \( \hat{x}(t_k-1) \). Constraint (24g) restricts that the difference between the state estimate given by the MHE (i.e., \( \hat{x}(t_k) \)) and the one given by the nonlinear observer (i.e., \( z(t_k) \)) is bounded in a region whose size depends on the difference between the current output measurement \( y(t_k) \) and the predicted output by the nonlinear observer (i.e., \( h(z(t_k)) \)). From Proposition 1, it is known that the difference between the estimate \( z(t_k) \) and the actual system state \( x(t_k) \) is bounded. However, the expression of the upper bound in (7) can not be used in the MHE formulation to characterize the region that contains the actual system state since the bound has dependence on \( x \) (recall the definition of \( e \) which is unknown). Fortunately, the measurement at \( t_k \) (i.e., \( y(t_k) \)) provides new information and the difference between \( y(t_k) \) and \( h(z(t_k)) \) (i.e., \( y(t_k) - h(z(t_k)) \)) can be used to measure the accuracy of estimate \( z(t_k) \). For example, when \( |y(t_k) - h(z(t_k))| \) is small, it implies that \( z(t_k) \) is very close to \( x(t_k) \) and a small region around \( z(t_k) \) can include the actual system state \( x \); when \( |y(t_k) - h(z(t_k))| \) increases, it implies that the accuracy of \( z(t_k) \) decreases and \( \hat{x}(t_k) \) should be restricted in a larger region around \( z(t_k) \) in order to include \( x(t_k) \) in the region.

In optimization problem (24), \( x_0 \in \mathbb{X} \) and output \( y \) sampled at time instants \( \{t_k\geq 0\} \). If the MHE scheme (24) is designed based on a deterministic nonlinear observer that satisfies condition (4) and initialized with \( \hat{x}_0 \in \mathbb{X} \) (i.e., \( e(t_0) = \hat{x}_0 - x_0 \)), and if there exists a concave function \( g(|e|) \) such that:
\[
g(|e|) \geq \beta(|e|, \Delta) \tag{26}
\]
for all \( |e| \leq d \) and positive constants \( d_s < d, a \geq 1, b, \epsilon \) such that:
\[
d_s - a(g(d_s) + \gamma_z(\Delta) + \gamma_{\epsilon}(\Delta)) - b \theta_v \geq \epsilon, \tag{27}
\]
and if the parameter \( \kappa \) is picked such that:
\[
0 \leq \kappa \leq \min\{(a-1)/L_h, b\}, \tag{28}
\]
then the estimation error \( |e| = |\hat{x} - x| \) is a decreasing sequence and is ultimately bounded as follows:
\[
\lim_{t \to \infty} \sup_{t \geq 0} |e(t)| \leq d_{\min} \tag{29}
\]
with \( d_{\min} = \max\{|e(t+\Delta)| : |e(t)| \leq d_s \} \) for all \( |e(t_0)| \leq d \) and \( x, z \in \mathbb{X} \).

**Proof:** First of all, it is noted that \( z(t_k) \) is a feasible solution to the MHE optimization problem (24) which implies that the optimization problem (24) is always feasible. In the remainder of the proof, it will be proved that the sequence of the values of the estimation error \( |e(t_0)| \) is a decreasing sequence if \( |e| \geq d_s \); and if \( |e| < d_s \), it will remain to be smaller than \( d_{\min} \) for ever. Specifically, the focus will be first on the evolution of \( |e| \) from \( t_{k-1} \) to \( t_k \) and then will be extended to the general case.

From the formulation of MHE (24), it can be written that:
\[
|\hat{x}(t_k) - z(t_k)| \leq \kappa|y(t_k) - h(z(t_k))|. \tag{30}
\]
From the Lipschitz property of \( h \), it is obtained that:
\[
|\hat{x}(t_k) - z(t_k)| \leq \kappa L_h|x(t_k) - z(t_k)| + \kappa \theta_v \tag{31}
\]
where $L_h$ the Lipschitz constant of $h$. Using the triangle inequality, it can be written that:

$$|\hat{x}(t_k) - x(t_k)| \leq |\hat{x}(t_k) - z(t_k)| + |z(t_k) - x(t_k)|. \quad (32)$$

From (31) and (32), it is obtained that

$$|e(t_k)| \leq (1 + \kappa L_h) |x(t_k) - z(t_k)| + \kappa \theta_v. \quad (33)$$

From Proposition 1 and (33), it can be obtained that:

$$|e(t_k)| \leq (1 + \kappa L_h) (\beta |e(t_k-1)| + |t - t_k-1|)$$

for $t \in [t_k-1, t_k]$. If condition (26) is satisfied, from inequality (34), it is obtained that:

$$|e(t_k)| \leq (1 + \kappa L_h) (g|e(t_k-1)| + \gamma_z (t - t_k-1) + \gamma_e (t - t_k-1)) + \kappa \theta_v \quad (35)$$

for $|e| \leq d$. If there exists a constant $0 < d_s < d$ satisfying (27) and $\kappa$ is picked following (28), then (27) holds for all $d_s \leq |e| \leq d$ taking into account that $g(\cdot)$ is a concave function, that is:

$$|e| - (1 + \kappa L_h) (g|e| + \gamma_z (\Delta) + \gamma_e (\Delta)) - \kappa \theta_v \geq \epsilon \quad (36)$$

for all $d_s \leq |e| \leq d$. From (35) and (36), it can be obtained that

$$|e(t_k)| \leq |e(t_k-1)| - \epsilon \quad (37)$$

for $d_s \leq |e| \leq d$. If $|e| \geq d_s$ for all the time from $t_0$ to $t_k$, using (37) recursively from $t = t_0$ to $t = t_k$, it can be obtained that:

$$|e(t_k)| \leq |e(t_0)| - k \epsilon \quad (38)$$

for all $d_s \leq |e(t_k)| \leq d$. This implies that $|e|$ decreases every sampling time and will become smaller than $d_s$ in finite steps. Once $|e| < d_s$, it will remain to satisfy $|e(t)| \leq d_{\text{min}}$ which is ensured by the definition of $d_{\text{min}}$. This implies that:

$$\lim_{t \to \infty} \sup_{t \in \mathbb{R}} |e(t)| \leq d_{\text{min}}. \quad (39)$$

This proves Theorem 1.

V. APPLICATION TO A GAS-PHASE BATCH REACTOR

Consider an isothermal gas-phase reactor in which a reversible reaction in the form $2A \rightleftharpoons B$ takes place. An initial amount of $A$ and $B$ are fed into the reactor. It is assumed that the ideal gas law holds and the reactor is well-mixed. Based on mass balance, the dynamics of partial pressures of $A$ and $B$, $P_A$ and $P_B$, can be described as follows [9]:

$$\frac{dP_A}{dt} = -2k_1 P_A^2 + 2k_2 P_B \quad (40)$$

$$\frac{dP_B}{dt} = k_1 P_A^2 - k_2 P_B$$

where $k_1 = 0.16 \text{ min}^{-1} \text{ atm}^{-1}$ and $k_2 = 0.64 \text{ min}^{-1}$ are the forward and reverse reaction rate constants. The initial composition of the process is $x_0 = [P_{A0}, P_{B0}]^T = [7 \text{ atm}, 1 \text{ atm}]^T$. It is assumed that the measurable output is the total pressure in the reactor; that is, $P = P_A + P_B$.

The measurements of the total pressure is subject to bounded measurement noise. Bounded random noise is also added to the right hand side of (40) to simulate bounded process model uncertainties. Specifically, the noise in the measurements of $P$ is generated as normal distributed values with zero mean and standard deviation $2$ (i.e., $v \sim \mathcal{N}(0, 2^2)$) but the values are restricted to be in the interval $[-0.8, 0.8]$. The disturbances in the dynamics of $P_A$ and $P_B$ are generated as $w \sim \mathcal{N}(0, 5^2)$ with the values restricted in $[-0.2, 0.2]$. We also assume that uncertainties are present in $k_1$ and $k_2$ of Eq. 40. The actual values used to evaluate the actual system trajectory are $1.1k_1$ and $1.1k_2$ (10% uncertainty) which are different from the values $(k_1$ and $k_2$) used in the design of observers.

For process (40), a Luenberger-like nonlinear observer can be designed as follows:

$$\frac{d\hat{P}_A}{dt} = -2k_1 \hat{P}_A^2 + 2k_2 \hat{P}_B + L_1 (P - \hat{P}_A - \hat{P}_B) \quad (41)$$

$$\frac{d\hat{P}_B}{dt} = k_1 \hat{P}_A^2 - k_2 \hat{P}_B + L_2 (P - \hat{P}_A - \hat{P}_B)$$

with $L_1 = L_2 = 5$. This observer is able to asymptotically track the nominal process state with continuous total pressure measurements.

We design the classical MHE and the proposed MHE based on the following weighting matrices:

$$Q = \begin{bmatrix} 25 & 0 \\ 0 & 25 \end{bmatrix}, R = 4.$$

Note that $Q$ and $R$ are chosen to approximate the distributions of the noise. Both MHE schemes use the same sampling time $\Delta = 0.05 \text{ min}$. In the proposed MHE scheme, $\kappa$ is tuned to be $0.4$. In the remainder of this section, the nonlinear observer (41) will also be assumed to be applied with measurements sampled every $\Delta$. In order to compare the performance of the different observers, the following performance index will be used:

$$J = \sum_{k=0}^{k_f} (\hat{x}(t_k) - x(t_k))^2 \quad (42)$$

where $t_0 = 0$ is the initial simulation time and $t_f = 2.5 \text{ min}$ is the end of simulation time, $x = [P_A, P_B]^T, S = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is a weighting matrix. The values of the weights in $S$ is chosen to account for the different ranges of numerical values for each state to make sure that different states contribute roughly equally to the value of $J$. In the design of the MHE schemes, the same extended Kalman filter is used to approximate the arrival cost.

First, the performance of the three observers are compared with the estimation horizon of the proposed and the classical MHE schemes being $N = 8$. Figures 1 and 2 show the results given by the two MHE schemes with a horizon $N = 8$ and the results given by the nonlinear observer (41). From these figures, we see that the two MHE schemes as well as the nonlinear observer are all able to track the evolution of
the process pressures. Specifically, the performance given by nonlinear observer (41), the classical MHE, and the proposed MHE are 60.24, 70.21 and 51.14. From these values, it can be seen that the proposed MHE gives the best performance.

Next, a set of simulations is carried out to study the boundedness of estimation errors of the proposed MHE scheme. In this set of simulations, the initial guess of the process is chosen to be $$\hat{x}_0 = [6.8 \text{ atm}, 1.2 \text{ atm}]^T$$ which is close to the actual initial state. The selection of the initial guess for the observers is to eliminate the effects of error in initial state guess and to focus on the ultimate boundedness of the estimation errors. In the simulations, the estimation horizon is $$N = 6$$ for both MHE schemes and the simulation duration is 5 min (i.e., 100 estimates in each run). Figure 3 shows the simulation results. Figures 3(a)-(b) show the distribution of estimation errors of $$P_A$$ and $$P_B$$ in the state space. From these figures, it can be seen that the estimation errors of $$P_A$$ and $$P_B$$ are bounded in a much smaller region when the proposed MHE scheme is used compared with the region given by the classical MHE scheme. Figures 3(c)-(d) show the absolute values of the estimation errors under the classical and the proposed MHE schemes, respectively. From the two figures, it can be seen that the maximum absolute value of the estimation error given by the proposed MHE scheme is 1.27 which is much smaller than the corresponding value 2.12 given by the classical MHE scheme. The results of this set of simulations demonstrate that boundedness of the estimation errors can be achieved by taking advantage the robustness of the deterministic observer (41) appropriately like in the proposed MHE scheme.

REFERENCES