Stability of a Nonlinear Moving Horizon Estimator with Pre-Estimation

Rata Suwantong, Sylvain Bertrand, Didier Dumur, Dominique Beauvois

Stability of a Nonlinear Moving Horizon Estimator with Pre-Estimation

Rata Suwantong[1], Sylvain Bertrand[1], Didier Dumur[2] and Dominique Beauvois[2]

Abstract— In this paper, a Moving Horizon Estimator with pre-estimation (MHE-PE) is proposed for discrete-time nonlinear systems under bounded noise. While the classical Moving Horizon Estimator (MHE) compensates for model errors by estimating the process noise sequence over the horizon via optimization, the MHE-PE does it using an auxiliary estimator. The MHE-PE is shown to require significantly less computation time compared to the MHE, while providing the same order of magnitude of estimation errors. The stability of the estimation errors of the MHE-PE is also proven and an upper bound on its estimation errors is derived. Performances of the MHE-PE is illustrated via a simulation example of pressure estimation in a gas-phase reversible reaction.

I. INTRODUCTION

A Moving Horizon Estimator (MHE) computes an estimate at the current instant by solving an optimization problem based on information from a fixed-number of latest measurements collected over a finite horizon. In this problem, the cost function to be minimized is traditionally described by the norm of the difference between real and predicted measurements over the horizon, a norm of the process noise over the horizon and a norm of the difference between an estimate at the beginning of the horizon and an a priori one. The latter term of the cost function is usually referred to as an "arrival cost" [5][6][10]. Once a new measurement is available, the oldest measurement is discarded and the horizon is moved forward. The main advantages of the MHE are that it allows to handle nonlinear systems without linearisation and to incorporate constraints directly during the optimization. The MHE has also been shown to be more robust against poor initialization than the Extended Kalman Filter (EKF) [5], the Unscented Kalman Filter (UKF) and the particle filter [9]. The convergence of the estimation errors of the MHE has been demonstrated in [2] provided that a quadratic arrival cost is adopted and its weight matrix is adequately chosen.

However, the computation time of the MHE can be very large due to its high number of optimization parameters which are not only the estimate at the beginning of the horizon but also the process noise sequence over the horizon. To decrease this computation time, fast optimization techniques for the MHE have been proposed in [3][10]. Another strategy to reduce computation time without changing the

optimization method consists in replacing the state equation in the MHE by an auxiliary estimator, called pre-estimating estimator. This pre-estimating estimator allows the MHE-PE to compensate for model errors without searching for the optimal process noise sequence over the horizon via optimization like the MHE. Such strategy has been proposed in [8] for discrete-time linear systems where a deterministic observer is chosen as pre-estimating estimator. It has been shown to be robust against model errors and require smaller computation time than the MHE since the process noise sequence is no longer included in the optimization parameters.

In this paper, an MHE with pre-estimation (MHE-PE) is proposed for discrete-time nonlinear systems under bounded noise. The novelties of this paper compared to [8] are the extension of the method to nonlinear systems and the possibility to use a bounded-error estimator instead of a deterministic observer in the pre-estimation part. The stability of the proposed MHE-PE is also analysed and an upper bound on the estimation errors is derived. Performance of the MHE-PE is illustrated via numerical simulations in terms of accuracy of the estimates and computation time. Note that an attempt to incorporate an observer in the MHE strategy for nonlinear systems under bounded noise has been proposed in [6]. However, in their strategy, the observer is only used to impose an additional constraint in the MHE and the process noise sequence is still included in the optimization parameters which induces large computation time.

The paper is organized as follows. In section II, some notations, definitions and general assumptions are introduced regarding the pre-estimating estimator and the MHE-PE formulation. In section III, an upper bound on the estimation errors provided by the pre-estimator is derived. In section IV, the stability of the estimation error dynamics of the MHE-PE is studied. In section V, the numerical studies are presented and the paper ends with concluding remarks.

II. PRELIMINARIES

A. Notations and Definitions

$\|v\|$ denotes the Euclidean norm of a vector v. A function $f(x)$ is said to be locally Lipschitz with respect to its argument x if there exists a positive 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 $\phi : [0,a) \rightarrow [0,\infty)$ is said to belong to class K if it is strictly increasing and satisfies $\phi(0) = 0$. A function $\beta(r,s)$ is said to be a class KL-function if, for each fixed s, $\beta(r,s)$ is a class K function with respect to r and, for each fixed r, $\beta(r,s)$ is decreasing with respect to s and $\beta(r,s) \rightarrow 0$ as $s \rightarrow \infty$.

[1]R. Suwantong and S. Bertrand are with the Department of System Design and Performance Evaluation, ONERA-The French Aerospace-lab, F-91761 Palaiseau, France rata.suwantong@onera.fr sylvain.bertrand@onera.fr

[2]D. Dumur and D. Beauvois are with the Department of Automatic Control, E3S, SUPELEC, 91192 Gif-sur-Yvette, France didier.dumur@supelec.fr dominique.beauvois@supelec.fr

B. System, Estimator and Associated Assumptions

Consider a nonlinear discrete-time system with additive noise of the form

$$\begin{aligned} x_{t+1} &= f(x_t) + w_t \\ y_t &= h(x_t) + v_t \end{aligned} \quad (1)$$

where $x_t \in \mathbb{X} \subset \mathbb{R}^n$ is the state, $y_t \in \mathbb{Y} \subset \mathbb{R}^m$ is the measurement, $w_t \in \mathbb{W} \subset \mathbb{R}^n$ is the process noise, $v_t \in \mathbb{V} \subset \mathbb{R}^m$ is the measurement noise and $t \in \mathbb{N}$ is the time index. The problem addressed in this paper consists in the estimation of the state vector x_t from the measurements $y_k (k \leq t)$. The system (1) will be referred to as the *real system*. The following assumptions are assumed to be satisfied:

(A1) $\mathbb{X}$ is a convex compact set, $\mathbb{W}$ and $\mathbb{V}$ are compact sets with $0 \in \mathbb{W}$ and $0 \in \mathbb{V}$. Define

$$r_w \triangleq \max_{w \in \mathbb{W}} \|w\|, \quad r_v \triangleq \max_{v \in \mathbb{V}} \|v\| \quad (2)$$

(A2) The initial state x_0 is such that, for any possible sequence of process noise $\{w_t\}$, the system trajectory $\{x_t\}$ lies in the convex compact set $\mathbb{X}$, $\forall t$

(A3) f and h are $\mathscr{C}^2$ functions with respect to x on $\mathbb{X}$. Consequently, f and h are also locally Lipschitz. Define their Lipschitz constants as L_f^x and L_h^x respectively

Define the *nominal system* as the "noise-free" part of the real system and denote $^nx_t \in \mathbb{X}$ the nominal state and $^ny_t \in \mathbb{R}^m$ the noise free output. The nominal system is defined as:

$$\begin{aligned} ^nx_{t+1} &= f(^nx_t) \\ ^ny_t &= h(^nx_t) \end{aligned} \quad (3)$$

Let us define an estimator map $g(\cdot, \cdot) \colon \mathbb{Z} \times \mathbb{R}^m \to \mathbb{Z}$, with $\mathbb{Z} \subset \mathbb{R}^n$ a convex compact set, by:

$$\mathfrak{z}_{t+1} = g(\mathfrak{z}_t, \gamma_t) \quad (4)$$

where $\mathfrak{z} \in \mathbb{Z}$ is the estimate given by g and $\gamma_t \in \mathbb{R}^m$ is the input of g. When y_t is used as input ($\gamma_t = y_t$), i.e. when the estimator is applied to the real system (1), the state $\mathfrak{z}_t$ will be denoted as z_t. In the same way, when ^ny_t is used as input ($\gamma_t = {^ny_t}$), i.e. when the estimator is applied to the nominal system (3), the state $\mathfrak{z}_t$ will be denoted as ^nz_t.

The estimator g is assumed to satisfy:

(A4) g is locally Lipschitz with respect to its arguments, with the associated Lipschitz constants L_g^z and L_g^y

(A5) $\forall z_0 \in \mathbb{Z}$, $\forall x_0 \in \mathbb{X}$, there exists a class K function ψ such that $\|{^nz_t} - {^nx_t}\|^2 \leq \psi(\|{^nz_0} - {^nx_0}\|^2)$, $\forall t$, i.e. the estimation error in a noise-free case is equal to zero if g is initialized at the real state, or

(A5a) g provides bounded estimation errors under bounded noise

An example of g that verifies (A5) is a discrete-time deterministic observer, i.e. g is designed such that $\lim_{t \to \infty} {^nz_t} = {^nx_t}$, $\forall {^nx_0} \in \mathbb{X}$ and $\forall {^nz_0} \in \mathbb{Z}$. In this case, there exists a KL-function β such that

$$\|{^nz_t} - {^nx_t}\| \leq \beta(\|{^nz_0} - {^nx_0}\|, t) \quad (5)$$

In section III, an upper bound on the estimation errors provided by g satisfying (A4) and (A5) will be derived.

C. Formulation of the MHE-PE

Assume that the state vector x_t has to be estimated at instant $t \geq N$ using the latest $N+1$ measurements collected within the "sliding horizon" $[t - N, t]$. Let us denote $z_{t-N,t}$ an estimate of x_{t-N} computed at instant t, $y_{t-N}^t = (y_{t-N}, y_{t-N+1}, \ldots, y_t)$ the measurement collection over the horizon and $\bar{z}_{t-N,t}^o$, an a priori value of $z_{t-N,t}$.

The MHE-PE is formulated as follows:

$$\hat{z}_{t-N,t}^o = \arg \min_{z_{t-N,t} \in \mathbb{Z}} \mathbf{J}_t(z_{t-N,t}, \bar{z}_{t-N,t}^o, y_{t-N}^t) \quad (6a)$$

$$\mathbf{J}_t = \mu \|z_{t-N,t} - \bar{z}_{t-N,t}^o\|^2 + \sum_{i=0}^N \|y_{t-N+i} - h(z_{t-N+i,t})\|^2 \quad (6b)$$

$$z_{t-N+i+1,t} = g(z_{t-N+i,t}, y_{t-N+i}), \ \forall i \in [0, N-1] \quad (6c)$$

where μ is a positive scalar representing the confidence in the a priori value $\bar{z}_{t-N,t}^o$ with respect to the observation model. The estimate $\hat{z}_{t,t}^\circ$ of x_t at time t provided by the MHE-PE is computed using (6c) for $i = 1, \ldots, N$ from $\hat{z}_{t-N,t}^o$, and $\bar{z}_{t-N,t}^o$ is determined from $\bar{z}_{t-N,t}^o \triangleq f(\hat{z}_{t-N-1,t-1}^\circ)$.

Note that in the MHE-PE the evolution of the state at *each instant* over the horizon is subject to the equation of the estimator g instead of the state equation as in the classical MHE. In fact, the MHE-PE locally implements the estimator g which is re-initialized at each instant at the beginning of the horizon by solving the optimization problem (6).

III. UPPER BOUND ON THE ESTIMATION ERRORS OF AN ESTIMATOR VERIFYING (A4) AND (A5)

In this section, an upper bound on the estimation errors provided by an estimator g verifying (A4) and (A5) is derived in case of bounded noise for system (1). The work is inspired by the method proposed in [6] where an upper bound on the estimation errors of a nonlinear continuous-time deterministic observer is derived.

Consider the real system starting from the initial state x_t evolving as in (1) and the nominal system starting from the same initial condition $^nx_t = x_t$ evolving as in (3). $\forall i \in \mathbb{N}^*$, denote z_{t+i} the estimate provided by the estimator g receiving real (noisy) measurements y_t^{t+i} starting from the initial estimate z_t. Denote $^nz_{t+i}$ the estimate provided by g receiving noise-free measurements $^ny_t^{t+i}$ and starting from the same initial estimate $^nz_t = z_t$.

We would like to find an upper bound on the square norm of the estimation error $\|z_{t+i} - x_{t+i}\|^2$ at time $t+i$. Thanks to the triangle inequality, we remark that

$$\begin{aligned} \|z_{t+i} - x_{t+i}\|^2 \leq & \ 3\|z_{t+i} - {^nz_{t+i}}\|^2 + 3\|{^nz_{t+i}} - {^nx_{t+i}}\|^2 \\ & + 3\|{^nx_{t+i}} - x_{t+i}\|^2 \end{aligned} \quad (7)$$

The second term of the r.h.s. of (7) verifies $\|{^nz_{t+i}} - {^nx_{t+i}}\|^2 \leq \psi(\|{^nz_t} - {^nx_t}\|^2) = \psi(\|z_t - x_t\|^2)$. For the third term of the r.h.s. of (7), using (1), (2), (3), assumption (A3) and the triangle inequality, the following propositions are derived by recurrence, $\forall i \in \mathbb{N}^*$.

Proposition 3.1: The square norm of the difference at $t+i$ between the real state x_{t+i} of (1) starting from x_t, and the nominal state $^n x_{t+i}$ of (3) starting from $^n x_t = x_t$ is bounded as $\|x_{t+i} - ^n x_{t+i}\|^2 \leq (\sum_{j=0}^{i-1} \{2(L_f^x)^2\}^j) 2 r_w^2$

Proposition 3.2: The square norm of the difference at $t+i$ between the real measurement y_{t+i} of (1) starting from x_t and the nominal measurement $^n y_{t+i}$ of the nominal state of (3) starting from $^n x_t = x_t$ is bounded as

$$\|y_{t+i} - ^n y_{t+i}\|^2 \leq 2(L_h^x)^2 (\sum_{j=0}^{i-1} \{2(L_f^x)^2\}^j) 2 r_w^2 + 2 r_v^2$$

For the first term of the r.h.s. of (7), by using (1), (2), (3), assumptions (A3) and (A4) along with the triangle inequality and by defining $\alpha_{t+j} = \sum_{l=0}^{j-1} \{2(L_f^x)^2\}^l$, $\forall j \in \mathbb{N}^+$ and $\alpha_t = 1$, the following proposition is derived.

Proposition 3.3: Consider system (1) starting from x_t and system (3) starting from $^n x_t = x_t$. $\forall i \in \mathbb{N}^*$, consider at $t+i$ the estimate z_{t+i} of the real state x_{t+i} given by g using y_t^{t+i} initialized at z_t. $\forall i \in \mathbb{N}^*$, consider at $t+i$ the estimate $^n z_{t+i}$ of the nominal state $^n x_{t+i}$ given by g using $^n y_t^{t+i}$ initialized at $^n z_t = z_t$. We have:

$$\|z_{t+i} - ^n z_{t+i}\|^2 \leq 2(L_g^y)^2 \times \dots$$
$$\left[4(L_h^x)^2 r_w^2 \sum_{j=0}^{i-1} \left(\{2(L_g^z)^2\}^{i-1-j} \alpha_{t+j} \right) + 2 r_v^2 \sum_{j=0}^{i-1} \{2(L_g^z)^2\}^j \right]$$

Proof: Let us start by considering when $i = 1$. Since $^n z_t = z_t$, it yields

$$\begin{aligned} \|z_{t+1} - ^n z_{t+1}\|^2 &= \|g(z_t, y_t) - g(^n z_t, ^n y_t)\|^2 \\ &\leq 2(L_g^y)^2 \|y_t - ^n y_t\|^2 \end{aligned}$$

Using proposition 3.2, the proposition 3.3 holds for $i = 1$. Now suppose that the proposition holds for $i \in \mathbb{N}^*$, let us prove that it is also the case for $i+1$.

$$\|z_{t+i+1} - ^n z_{t+i+1}\|^2 = \|g(z_{t+i}, y_{t+i}) - g(^n z_{t+i}, ^n y_{t+i})\|^2$$
$$\leq 2(L_g^z)^2 \|z_{t+i} - ^n z_{t+i}\|^2 + 2(L_g^y)^2 \|y_t - ^n y_t\|^2$$
$$\leq 4(L_g^y)^2 \left[r_v^2 \sum_{j=0}^{i} \{2(L_g^z)^2\}^j + 2(L_h^x r_w)^2 \sum_{j=0}^{i} \{2(L_g^z)^2\}^{i-j} \alpha_{t+j} \right]$$

∎

Theorem 3.4: The square norm of the difference between the estimate z_{t+i} given by an estimator g verifying (A4) and (A5) receiving measurements y_t^{t+i} initialized at time t by z_t and the state x_{t+i} of the real system, $\forall i \in \mathbb{N}^*$, is bounded as:

$$\|z_{t+i} - x_{t+i}\|^2 \leq 3 \psi(\|z_t - x_t\|^2) + c_{w,i} r_w^2 + c_{v,i} r_v^2$$

where

$$c_{w,i} = 6 \sum_{j=0}^{i-1} \{2(L_f^x)^2\}^j + 24(L_g^y L_h^x)^2 \alpha_{t+j}$$

$$c_{v,i} = 12(L_g^y)^2 \sum_{j=0}^{i-1} \{2(L_g^z)^2\}^j \qquad (8)$$

Proof: The proof is straightforward by combining (8) and proposition 3.3 in (7), recalling that $^n z_t = z_t$ and $^n x_t = x_t$. ∎

Now that an upper bound on the estimation errors provided by an estimator verifying (A4) and (A5), applied to a system under bounded noise, has been determined, this result will be used to analyse the error dynamics of the MHE-PE when such pre-estimating estimator is chosen. The following study is also verified for g verifying (A5a) instead of (A5) with different values of associated constants.

IV. ERROR DYNAMICS OF THE MHE-PE

To determine the dynamics of the estimation error of the MHE-PE, upper and lower bounds on the optimal cost, noted by $\mathbf{J}_t^o \triangleq \mathbf{J}_t(\hat{z}_{t-N,t}^o, \bar{z}_{t-N,t}^o, y_{t-N}^t)$ where $\hat{z}_{t-N,t}^o$ is the optimal solution of the problem (6), are calculated. The approach is inspired by [1] where the error dynamics of the MHE (without pre-estimation) is derived.

A. Upper Bound on the Optimal Cost

Denote x_{t-N}, the state of (1) at instant $t-N$. By optimality of $\hat{z}_{t-N,t}^o$, we get $\mathbf{J}_t^o \leq \mathbf{J}_t(x_{t-N}, \bar{z}_{t-N,t}^o, y_{t-N}^t)$, i.e.

$$\mathbf{J}_t^o \leq \mu \|x_{t-N} - \bar{z}_{t-N,t}^o\|^2 + \sum_{i=0}^{N} \|y_{t-N+i} - h(z_{t-N+i,t})\|^2 \quad (9)$$

with $z_{t-N+i+1,t} = g(z_{t-N+i,t}, y_{t-N+i}), \forall i \in [0, N-1]$
$\qquad z_{t-N,t} = x_{t-N}$

Let us find an upper bound on the second term of the r.h.s. of (9). Using the triangle inequality, we have

$$\|y_{t-N+i} - h(z_{t-N+i,t})\|^2$$
$$\leq 2\|y_{t-N+i} - h(x_{t-N+i})\|^2 + 2\|h(x_{t-N+i}) - h(z_{t-N+i,t})\|^2$$
$$\leq 2 r_v^2 + 2(L_h^x)^2 \|x_{t-N+i} - z_{t-N+i,t}\|^2$$

Therefore,

$$\sum_{i=0}^{N} \|y_{t-N+i} - h(z_{t-N+i,t})\|^2 \leq$$
$$2 \sum_{i=0}^{N} r_v^2 + 2(L_h^x)^2 \sum_{i=0}^{N} \|x_{t-N+i} - z_{t-N+i,t}\|^2 \quad (10)$$

To bound the second term of the r.h.s. of (10), we use theorem 3.4 recalling that $\psi(\|z_{t-N,t} - x_{t-N}\|^2) = 0$ since $z_{t-N,t} = x_{t-N}$.

$$\sum_{i=0}^{N} \|y_{t-N+i} - h(z_{t-N+i,t})\|^2 \leq l_{w,N} r_w^2 + l_{v,N} r_v^2 \quad (11a)$$

$$l_{w,N} = 2(L_h^x)^2 \sum_{i=1}^{N} c_{w,i} \quad (11b)$$

$$l_{v,N} = 2(L_h^x)^2 \sum_{i=1}^{N} c_{v,i} + 2(N+1) \quad (11c)$$

where $c_{w,i}$ and $c_{v,i}$ are defined in (8).
Hence, an upper bound on the optimal cost $\mathbf{J}_t^\circ$ is given by [1]

$$\mathbf{J}_t^o \leq \mu \|x_{t-N} - \bar{z}_{t-N,t}^o\|^2 + c_N^2 \quad (12)$$

where $\qquad c_N^2 = l_{w,N} r_w^2 + l_{v,N} r_v^2 \quad (13)$

Now that we have an upper bound on the optimal cost, let us pursue with the computation of a lower bound.

[1]Note that for g satisfying (A5a), only the value of the constant c_N will change.

B. Lower Bound on the Optimal Cost

We start by determining a lower bound on any common cost $\mathbf{J}_t$ in (6), so $z_{t-N,t}$ will be replaced by z_{t-N} to refer to any value of the estimate at the beginning of the horizon in general. After that, a lower bound on the optimal cost $\mathbf{J}_t^\circ$ will be derived by replacing z_{t-N} by $\hat{z}_{t-N,t}^\circ$.

Let us define first the observation maps of the real system in (1) and of the pre-estimating estimator g initialized at $t-N$ by z_{t-N} receiving the real measurements y_{t-N}^t respectively by

$$F(x_{t-N}) \triangleq \begin{pmatrix} h(x_{t-N}) \\ h(f(x_{t-N})) \\ \vdots \\ h(f^N(x_{t-N})) \end{pmatrix} \quad (14)$$

and

$$G(z_{t-N}, y_{t-n}^{t-1}) \triangleq \begin{pmatrix} h(z_{t-N}) \\ h \circ g(z_{t-N}, y_{t-N}) \\ \vdots \\ h(g^N(z_{t-N}, y_{t-N}^{t-1})) \end{pmatrix} \quad (15)$$

where for $i \geq 1$

$$g^i(z_{t-N}, y_{t-N}^{t-N+i-1}) \triangleq \underbrace{g(g \dots g}_{i \text{ times}}(z_{t-N}, y_{t-N}), \dots, y_{t-N+i-1})$$

Therefore, the second term of the r.h.s. of (6b) becomes

$$\sum_{i=0}^{N} \|y_{t-N+i} - h(z_{t-N+i})\|^2 = \|y_{t-N}^t - G(z_{t-N}, y_{t-N}^{t-1})\|^2.$$

Denote $^n y_{t-N}^{t-1}$ the noise free measurements provided by system (3) initialized at time $t-N$ with $^n x_{t-N} = x_{t-N}$. Thanks to the triangle inequality, we deduce:

$$\begin{aligned}
\|y_{t-N}^t - G(z_{t-N}, y_{t-N}^{t-1})\|^2 &\geq \\
\frac{1}{4} \|G(z_{t-N}, {}^n y_{t-N}^{t-1}) &- G(x_{t-N}, {}^n y_{t-N}^{t-1})\|^2 \\
- \|y_{t-N}^t - F(x_{t-N})\|^2 &- \|F(x_{t-N}) - G(x_{t-N}, {}^n y_{t-N}^{t-1})\|^2 \\
- \|G(z_{t-N}, y_{t-N}^{t-1}) &- G(z_{t-N}, {}^n y_{t-N}^{t-1})\|^2
\end{aligned} \quad (16)$$

To calculate a lower bound on the first term of the r.h.s. of (16), consider the system of the estimator g in (4) as a nonlinear system as defined in [4] where measurements are considered as the input of the system. It is shown in [4] that a lower bound on the term of interest exists if the following assumptions are satisfied:

(A6) g is K-uniformly observable on $\mathbb{Z}$ with respect to all admissible measurements, i.e. $\exists N > 0$, $\forall (z', z'') \in \mathbb{Z}^2$, $\forall y \in \mathbb{Y}$, there exists a K-function $\phi(\cdot)$ such that

$$\phi(\|z' - z''\|^2) \leq \|G(z', y_{t-N}^{t-1}) - G(z'', y_{t-N}^{t-1})\|^2 \quad (17)$$

(A7) The observation map $G(\cdot, \cdot)$ has a finite sensitivity to the estimate, i.e. the K-function $\phi(\cdot)$ in (17) satisfies:

$$\delta = \inf_{(z_1, z_2) \in \mathbb{Z}^2, z_1 \neq z_2} \frac{\phi(\|z_1 - z_2\|^2)}{\|z_1 - z_2\|^2} > 0 \quad (18)$$

Thanks to (17) and (18), we obtain a lower bound on the first term of the r.h.s. of (16)

$$\|G(x_{t-N}, {}^n y_{t-N}^{t-1}) - G(z_{t-N}, {}^n y_{t-N}^{t-1})\|^2 \geq \delta \|z_{t-N} - x_{t-N}\|^2 \quad (19)$$

Consider now the second term of the r.h.s. of (16), similarly to [1], we have

$$\|y_{t-N}^t - F(x_{t-N})\|^2 \leq c_{Ale,N}^2 \quad (20)$$

$$c_{Ale,N} \triangleq \Delta_w \sqrt{N} r_w + \sqrt{N+1} r_v + \frac{\bar{k}}{2} \sqrt{\frac{N(N+1)(2N+1)}{6}} r_w^2$$

where Δ_w and $\bar{k}$ characterize the model sensitivity to state noise, see [1]. The third term of the r.h.s. of (16) is equal to zero since g satisfies (A5)[2] Now, consider the forth term of the r.h.s of (16)

$$\|G(z_{t-N}, y_{t-N}^{t-1}) - G(z_{t-N}, {}^n y_{t-N}^{t-1})\|^2 \leq (L_h^x)^2 \sum_{i=1}^{N} \mathcal{G}_i \quad (21)$$

where $\mathcal{G}_i \triangleq \|g^i(z_{t-N}, y_{t-N}^{t-N+i-1}) - g^i(z_{t-N}, {}^n y_{t-N}^{t-N+i-1})\|^2$. By recurrence, it can be shown that

$$\mathcal{G}_i \leq \underbrace{2(L_g^z)^2 \mathcal{G}_{i-1}}_{\text{for } i \geq 1} + 2(L_g^y)^2 \|y_{t-N+i-1} - {}^n y_{t-N+i-1}\|^2 \quad (22)$$

To calculate an upper bound on $\mathcal{G}_i$, consider the following proposition:

Proposition 4.1: The square norm of the difference at time $t-N+i$ between the estimate provided by the estimator g receiving the collection of real measurements $y_{t-N}^{t-N+i-1}$ and initialized at $t-N$ with z_{t-N} and the estimate provided by the estimator g initialized at $t-N$ with $^n z_{t-N} = z_{t-N}$ but receiving the collection of the noise-free measurements $^n y_{t-N}^{t-N+i-1}$ associated to the nominal system (3) initialized at $t-N$ with $^n x_{t-N} = x_{t-N}$, is bounded as

$$\mathcal{G}_i \leq 2 \left(L_g^y\right)^2 \times \dots$$

$$\left(\underbrace{2r_v^2 \sum_{j=0}^{i-1} \left(2\left(L_g^z\right)^2\right)^j}_{\text{for } i \geq 1} + \underbrace{2(L_h^x)^2 2r_w^2 \sum_{j=0}^{i-2} \left(2(L_f^x)^2 + 2(L_g^z)^2\right)^j}_{\text{for } i \geq 2} \right)$$

Proof: It is easy to verify that the proposition 4.1 holds for $i = 1$ and $i = 2$ using (22) and proposition 3.2, where $t-N$ is considered as the first instant of the window. Now suppose that the proposition 4.1 holds for i and prove that it is also the case for $i+1$. Equation (22) and proposition 4.1 give

$$\begin{aligned}
\mathcal{G}_{i+1} \leq\ & 2(L_g^y)^2 \times \\
& \left[2r_v^2 \sum_{j=1}^{i} \left(2(L_g^z)^2\right)^j + 4(L_h^x L_g^z)^2 2r_w^2 \sum_{j=0}^{i-2} \left(2(L_f^x)^2 + 2(L_g^z)^2\right)^j \right] \\
& + 2(L_g^y)^2 2(L_h^x)^2 \left(\sum_{j=0}^{i-1} \left(2(L_f^x)^2\right)^j \right) 2r_w^2 + 2(L_g^y)^2 2r_v^2
\end{aligned}$$

$$\begin{aligned}
\mathcal{G}_{i+1} \leq\ & 2(L_g^y)^2 \left(2r_v^2 \sum_{j=0}^{i} \left(2(L_g^z)^2\right)^j \right) + 2(L_h^x)^2 2(L_g^y)^2 \times \\
& \left[2(L_g^z)^2 \sum_{j=0}^{i-2} \left(2(L_f^x)^2 + 2(L_g^z)^2\right)^j + \sum_{j=0}^{i-1} \left(2(L_f^x)^2\right)^j \right] 2r_w^2
\end{aligned}$$

---

[2]This term is equal to a constant if g verifies (A5a) instead.

Using lemma 6.1 in appendix, proposition 4.1 holds for $i+1$, which completes the proof by recurrence. ∎

Reconsider (21). Proposition 4.1 leads to

$$\left\|G(z_{t-N}, y_{t-N}^{t-1}) - G(x_{t-N}, {}^n y_{t-N}^{t-1})\right\|^2 \le (L_h^x)^2 \lambda_N^2 \qquad (23)$$

where

$$\lambda_N^2 = 2(L_g^y)^2 \times \{2r_v^2 \sum_{i=1}^N \sum_{j=0}^{i-1} \left(2(L_g^z)^2\right)^j \qquad (24)$$

$$+ 2(L_h^x)^2 2r_w^2 \sum_{i=2}^N \sum_{j=0}^{i-2} \left(2(L_f^x)^2 + 2(L_g^z)^2\right)^j\}$$

Replacing (19), (20), and (23) in (16), we finally obtain a lower bound on the second term of the r.h.s. of (6b)[3]

$$\left\|y_{t-N}^t - G(z_{t-N}, y_{t-N}^t)\right\|^2 \ge$$
$$\frac{1}{4}\delta \|z_{t-N} - x_{t-N}\|^2 - (L_h^x)^2 \lambda_N^2 - c_{Ale,N}^2 \qquad (25)$$

Let us continue with a lower bound on the first term of (6b) with the general notation z_{t-N} introduced at the beginning of the section: $\mu\|z_{t-N} - \bar{z}_{t-N}^o\|^2$. First of all, we remark that

$$\left\|z_{t-N} - \bar{z}_{t-N,t}^\circ\right\|^2 \ge \frac{1}{2}\|x_{t-N} - z_{t-N}\|^2 - \left\|x_{t-N} - \bar{z}_{t-N,t}^\circ\right\|^2 \qquad (26)$$

Combine (25) to (26) and replace the general notation z_{t-N} by the solution $\hat{z}_{t-N,t}^\circ$. By defining $e_{t-N}^\circ \triangleq \hat{z}_{t-N,t}^\circ - x_{t-N}$, a lower bound on the optimal cost $\mathbf{J}_t^\circ$ is derived as follows:

$$\mathbf{J}_t^\circ \ge \left(\frac{1}{2}\mu + \frac{1}{4}\delta\right)\left\|e_{t-N}^\circ\right\|^2 - \mu\left\|x_{t-N} - \bar{z}_{t-N}^\circ\right\|^2$$
$$- (L_h^x)^2 \lambda_N^2 - c_{Ale,N}^2 \qquad (27)$$

Now that a lower bound on the optimal cost $\mathbf{J}_t^\circ$ has been calculated, it will be combined with the upper bound on $\mathbf{J}_t^\circ$ computed in the previous section to analyse the dynamics of the estimation error of the MHE-PE.

C. Dynamics of the Estimation Errors

Theorem 4.2: Consider a discrete-time nonlinear system verifying assumptions (A1)-(A3). If the pre-estimating estimator of the MHE-PE verifies (A4)-(A7), then the square norm of the estimation error of the MHE-PE is bounded as $\left\|e_{t-N}^\circ\right\|^2 \le \zeta_{t-N}$ where $\{\zeta_t\}$ is a sequence generated by

$$\zeta_0 = \beta_0 \qquad (28)$$
$$\zeta_t = \alpha_N \zeta_{t-1} + \beta_N, \quad \forall t \in \mathbb{N}^*$$

with

$$\alpha_N = \frac{8\mu(L_f^x)^2}{\mu + \frac{\delta}{2}} \qquad \beta_N = \frac{2}{\mu + \frac{\delta}{2}}(4\mu r_w^2 + \mathscr{C}_N^2)$$

$$\beta_0 = \frac{2}{\mu + \frac{\delta}{2}}(2\mu d_x^2 + \mathscr{C}_N^2)$$

where $\mathscr{C}_N^2 \triangleq (L_h^x)^2 \lambda_N^2 + c_{Ale,N}^2 + c_N^2$[4] and $d_x \triangleq \max_{(x,x')\in\mathbb{X}^2} \|x - x'\|$.

Moreover, if μ is selected s.t.

$$8\mu(L_f^x)^2/(\mu + \frac{\delta}{2}) < 1 \qquad (29)$$

[3]For g satisfying (A5a), another constant will be added on the r.h.s.
[4]$\mathscr{C}_N$ takes another constant value for g verifying (A5a).

The sequence $\{\zeta_t\}$ has the following properties:

(a) $\{\zeta_t\}$ converges exponentially to the asymptotic value $e_\infty^\circ(\mu) \triangleq \beta_N/(1 - \alpha_N)$
(b) if $\zeta_t > e_\infty^\circ(\mu)$, then $\zeta_t < \zeta_{t-1}$

Proof: Combine the upper bound on the optimal cost $\mathbf{J}_t^\circ$ in (12) with its lower bound in (27) to get:

$$\frac{1}{2}\left(\mu + \frac{1}{2}\delta\right)\left\|e_{t-N}^\circ\right\|^2 \le 2\mu\left\|x_{t-N} - \bar{z}_{t-N,t}^\circ\right\|^2 + \mathscr{C}_N^2$$

If $\bar{z}_{t-N,t}^\circ \triangleq f(\hat{z}_{t-N-1,t-1}^\circ)$, therefore

$$\left\|x_{t-N} - \bar{z}_{t-N,t}^\circ\right\|^2$$
$$= \left\|f(x_{t-N-1}) + w_{t-N-1} - f(\hat{z}_{t-N-1,t-1}^\circ)\right\|^2$$
$$\le 2(L_f^x)^2\left\|e_{t-N-1}^\circ\right\|^2 + 2r_w^2 \qquad (30)$$

Thanks to (30), we have

$$\frac{1}{2}\left(\mu + \frac{1}{2}\delta\right)\left\|e_{t-N}^\circ\right\|^2 \le 4\mu(L_f^x)^2\left\|e_{t-N-1}^\circ\right\|^2 + 4\mu r_w^2 + \mathscr{C}_N^2 \qquad (31)$$

Applying condition (29) to (31) completes the proof. ∎

Thanks to theorem 4.2, the convergence of the estimates provided by MHE-PE can be guaranteed by choosing an appropriate μ satisfying condition (29). This is easy to do for any value of L_f^x, once δ is calculated using (18) assuming that the reachable values of the estimates can be considered as equal to those of the real system.

In the case of a noise-free system ($r_w = r_v = 0$), the asymptotic convergence of the MHE-PE is guaranteed by corollary 4.3.

Corollary 4.3: Suppose that assumptions (A1)-(A7) hold. If $r_w = r_v = 0$, $\left\|e_{t-N}^\circ\right\|^2 \le \alpha_N^{t-N}\beta_0$, $\forall t \ge N$. Moreover, if μ satisfies (29), $\lim_{t\to\infty}\left\|e_{t-N}^\circ\right\|^2 = 0$.

V. NUMERICAL EXAMPLE

The performances of the MHE-PE and of a classical MHE whose formulation can be found in [7] are compared through a pressure estimation problem of a gas-phase, reversible reaction as defined in [5]. The state is defined as the partial pressures of the system $x_t = \begin{pmatrix} x_{1,t} & x_{2,t} \end{pmatrix}^T$. Define $\bar{k}_g = 0.16$ $(atm \cdot s)^{-1}$ and the sampling time $T_s = 0.1$ s. The dynamics of the real system is:

$$x_{t+1} = f(x_t) + w_t = \begin{pmatrix} \dfrac{x_{1,t}}{2\bar{k}_g T_s x_{1,t} + 1} \\ x_{2,t} + \dfrac{\bar{k}_g T_s x_{1,t}^2}{2\bar{k}_g T_s x_{1,t} + 1} \end{pmatrix} + w_t$$

$$y_t = \begin{pmatrix} 1 & 1 \end{pmatrix} x_t + v_t, \quad x_0 = \begin{pmatrix} 5 & 1 \end{pmatrix}^T \qquad (32)$$

The noise w_t and v_t are supposed to be independent random variables, uniformly distributed in the intervals $[-0.06, 0.06] \times [-0.3, 0.3]$ and $[-0.3, 0.3]$, respectively. 100 Monte Carlo simulations of the system have been simulated for $t \in [0, 10]$ s using MATLAB on a standard PC. The a priori value $\bar{z}_0^\circ = \begin{pmatrix} 2 & 4.5 \end{pmatrix}^T$ and the constraint $\mathbb{Z} = [0, 5]^2$ are imposed to every MHE and MHE-PE for every run. The
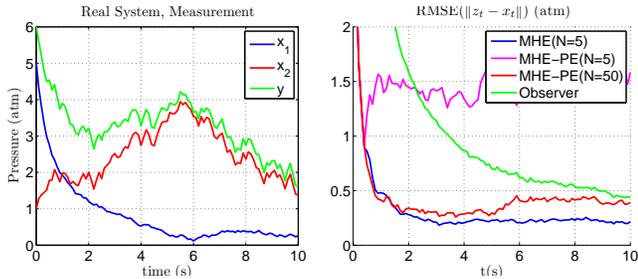
Fig. 1: (Left) Example of the real system state and measurement from one Monte Carlo run. (Right) RMSE given by each tested estimator

process noise estimates given by the MHE are imposed to be in the interval $[-0.3, 0.3]^2$. A deterministic observer for the noise-free system associated to system (32) is

$$z_{t+1} = g(z_t, y_t) = f(z_t) + L(y_t - \begin{pmatrix} 1 & 1 \end{pmatrix} z_t) \qquad (33)$$

where $L = \begin{pmatrix} 0.0026 & 0.7046 \end{pmatrix}^T$ has been computed using proposition 3 in [11]. The values of $L_f^x = 1$, $\delta = 1$, $\mu = 5 \cdot 10^{-4}$ are chosen. This value of μ gives $\alpha = 0.008$. $N = 5$ is chosen for both MHE and MHE-PE and $N = 50$ is chosen in addition for MHE-PE. Results for MHE with $N = 50$ are not included due to its too large computation time. All available measurements are used for $t < N$ for every MHE.

The accuracy of the estimates are studied in terms of Root Mean Square Error (RMSE) and Asymptotic Root Mean Square Error (ARMSE) as defined in [1]. The evaluation batch size for ARMSE calculation is set to 50. The RMSEs provided by the MHE, the MHE-PE and the nonlinear deterministic observer defined in (33) are shown in figure 1. To finish, the ARMSE and the mean computation times per iteration are presented in table 1.

MHE	N=5		N=50	
	ARMSE	mean time/ iteration (s)	ARMSE	mean time/ iteration (s)
MHE-PE	1.4971	0.0526	0.3997	0.0822
MHE	0.2231	0.4831	×	×

Table 1: ARMSE and mean computation time per iteration of MHE-PE for N=5 and N=50 and MHE for N=5.

It can be observed from figure 1 that the deterministic observer gives high RMSE and converges slowly and that a too small horizon can lead MHE-PE to have large estimation errors. However, increasing the horizon size does not affect much on MHE-PE's computation time since the number of the optimization parameters does not change. Moreover, when N is sufficiently high, MHE-PE gives estimation errors closed to those of MHE but uses much smaller computation time. According to table 1, MHE-PE can be used in real-time applications for this example since its computation time per iteration is lower than the sampling period.

To conclude, MHE-PE leads to a good trade-off between accuracy and small computation time making it an interesting alternative MHE strategy for discrete-time nonlinear systems under bounded noise, for which an estimator verifying (A4)-(A7) can be designed.

VI. CONCLUSIONS

A Moving Horizon Estimator with Pre-Estimation (MHE-PE) has been proposed for a nonlinear discrete-time system under bounded noise. The stability of the estimation error of the MHE-PE has been demonstrated and its upper bound has been derived. It has been shown via numerical simulations that the performance of the proposed MHE-PE offers a good trade-off between accuracy of the estimates and small computation time. Therefore, the MHE-PE can be considered as an alternative estimator to save computation time which opens the possibilities of real-time applications of Moving Horizon strategies.

APPENDIX

Lemma 6.1: If $a > 0$ and $b > 0$, for $i \geq 2$ we have

$$a \sum_{j=0}^{i-2}(a+b)^j + \sum_{j=0}^{i-1} b^j \leq \sum_{j=0}^{i-1}(a+b)^j \qquad (34)$$

Proof: It is evident that the lemma holds for $i = 2$. Now suppose that it holds for i, it holds for $t = i+1 \Leftrightarrow$

$$a \sum_{j=0}^{i-2}(a+b)^j + a(a+b)^{i-1} + \sum_{j=0}^{i-1} b^j + b^i \leq \sum_{j=0}^{i-1}(a+b)^j + (a+b)^i$$

Thanks to (34), there is only $a(a+b)^{i-1} + b^i \leq (a+b)^i$ left to be verified. Since $a > 0$ and $b > 0$, we have

$$\left(\frac{b}{a+b}\right)^i \leq \frac{b}{a+b} \Rightarrow \frac{b^i}{(a+b)^{i-1}} \leq b \Rightarrow a + \frac{b^i}{(a+b)^{i-1}} \leq a+b$$
$$\Rightarrow a(a+b)^{i-1} + b^i \leq (a+b)^i \qquad \blacksquare$$

ACKNOWLEDGMENT

REFERENCES

[1] A. Alessandri, M. Baglietto, and G. Battistelli, "Moving horizon state estimation for nonlinear discrete-time systems: New stability results and approximation schemes," *Automatica*, vol. 44, 2008.
[2] A. Alessandri, M. Baglietto, G. Battistelli, and V. Zavala, "Advances in moving horizon estimation for nonlinear systems," in *49th IEEE Conference on Decision and Control*, 2010.
[3] H. Ferreau, T. Kraus, M. Vukov, W. Saeys, and M. Diehl, "High-speed moving horizon estimation based on automatic code generation," in *51th IEEE Conference on Decision and Control*, 2012.
[4] S. Hanba, "Further results on the uniform observability of discrete-time nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, 2010.
[5] E. L. Haseltine and J. B. Rawlings, "Critical evaluation of extended Kalman filtering and moving-horizon estimation," *Industrial & Engineering Chemistry Research*, vol. 44, no. 8, 2005.
[6] J. Liu, "Moving horizon state estimation for nonlinear systems with bounded uncertainties," *Chemical Engineering Science*, vol. 93, 2013.
[7] C. Rao, J. Rawlings, and D. Mayne, "Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations," *IEEE Transactions on Automatic Control*, vol. 48, no. 2, 2003.
[8] D. Sui, T. Johansen, and L. Feng, "Linear moving horizon estimation with pre-estimating observer," *IEEE Transactions on Automatic Control*, vol. 55, no. 10, 2010.
[9] R. Suwantong, P. Bui Quang, D. Beauvois, D. Dumur, and S. Bertrand, "Robustness analysis of a moving horizon estimator for space debris tracking during atmospheric reentry," in *52th IEEE Conference on Decision and Control*, 2013.
[10] V. Zavala, "A fast moving horizon estimation algorithm based on nonlinear programming sensitivity," *Journal of Process Control*, vol. 18, no. 9, 2008.
[11] W. Zhang, H. Su, F. Zhu, and D. Yue, "A note on observers for discrete-time Lipschitz nonlinear systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 59, no. 2, 2012.