Stability Analysis and Robustness Assessment of Deterministic and Stochastic Nonlinear Moving Horizon Estimators

Rata Suwantong, Sylvain Bertrand, Didier Dumur, Dominique Beauvois

Stability Analysis and Robustness Assessment of Deterministic and Stochastic Nonlinear Moving Horizon Estimators

Rata Suwantong, Sylvain Bertrand, Didier Dumur and Dominique Beauvois

Abstract—This paper proposes a discussion on the classification of the formulations of nonlinear Moving Horizon Estimators (MHE) of the literature into two categories: deterministic and stochastic. The stability of the dynamics of the estimation error is discussed for the MHEs in both frameworks. This paper also provides full explicit formulation of the stability conditions for the MHE in the deterministic framework, which were not given in the literature. Furthermore, robustness of MHE in both frameworks with respect to model errors is investigated through a simulation example of space object tracking. Comparison with other more classical estimators such as EKF, UKF and particle filter is also achieved.

I. INTRODUCTION

The 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. Its optimization parameters are generally the initial state at the begining of the horizon and the process noise sequence over the horizon. The MHE has been proven to be more robust against model errors, poor initialization and bad tuning compared to classical filters such as the Extended Kalman Filter, the Unscented Kalman Filter and the particle filters [1][2][3].

For discrete time autonomous nonlinear systems with additive process and measurement noises, formulations of the MHE exist in the literature in both deterministic and stochastic frameworks. The deterministic framework refers to the fact that the initial state at the begining of the horizon, the process noise and the measurement noise are considered to be unknown deterministic variables of unknown characteristics that take their values in known compact sets [4][5][6]. The stochastic framework refers to the case where these parameters are considered as random variables and assumptions on their statistics are made [2][7][8].

A major issue in nonlinear MHE is to provide stability guarantees of the estimation error dynamics. In the stochastic framework, some stability conditions have been provided in the literature [7] but they are difficult to deal with for practical implementation. On the other hand, more tractable conditions have been derived for MHE in the deterministic framework to ensure bounded convergence of the estimation error [5][6].

R. Suwantong is with the Geo-Informatics and Space Technology (GISTDA), 10210, Bangkok, Thailand rata@gistda.or.th

S. Bertrand is with ONERA-The French Aerospacelab, F-91761, Palaiseau, France sylvain.bertrand@onera.fr

D. Dumur and D. Beauvois are with the Department of Automatic Control, Ecole Supérieure d'Electricité, 91192, Gif-sur-Yvette, France didier.dumur@centralesupelec.fr dominique.beauvois@centralesupelec.fr

These stability conditions are based on mild conditions on the continuity of the state equation, the observability of the system and the finite sensitivity to the measurements which are easy to be verified in practice. Another condition concerns the weight term on the initial state at the beginning of the horizon which is used in the cost function of the MHE formulation. Explicit formulation of all the terms involved in this condition, and which are useful for practical implementation, has not been given in the literature. This paper therefore provides the formulations of all the terms used in the stability theorem of the MHE in the deterministic framework.

Another concern in state estimation is robustness of the developed estimators with respect to model errors. This paper hence proposes such a robustness assessment of the MHE in both frameworks through a simulation example of space debris tracking during atmospheric re-entry on Earth, which is an nonlinear estimation problem where high model errors are present.

This paper is organized as follows. In section II, the notations used in the paper, the system definition and a generic formulation of the MHE problem are provided. In section III, differences between the MHE in both frameworks are discussed in terms of assumptions, parameter tuning, and stability of the dynamics of the estimation error. In section IV, robustness against model errors of the MHE in both frameworks is studied through the proposed simulation example.

II. PRELIMINARIES

A. Notations

The Euclidean norm of a vector v is denoted by $\|v\|$. For a given matrix Q and vector v, we define $\|v\|_Q^2 = v^T Q v$. The maximum and minimum eigenvalue of a matrix Q will be respectively denoted by $\lambda_{\mathbf{Q}}^{max}$ and $\lambda_{\mathbf{Q}}^{min}$. 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 $\|f(x') - f(x'')\| \le L_f^x \|x' - x''\|$, for all x' and x'' in a given region of x and L_f^x is the associated Lipschitz constant. A continuous function $\phi : [0, a) \to [0, \infty)$ 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 $\beta(r, s) \to 0$ as $s \to \infty$.

B. Model of the System: State and Measurement Equations

The considered autonomous system is modeled by the nonlinear discrete-time equations:

$$x_{k+1} = f(x_k) + w_k, \quad y_k = h(x_k) + v_k \tag{1}$$

where $x_k \in \mathbb{R}^{n_x}$ is the state vector of dimension n_x, $y_k \in \mathbb{R}^{n_y}$ is the measurement vector, $w_k \in \mathbb{R}^{n_x}$ is the process noise, $v_k \in \mathbb{R}^{n_y}$ is the measurement noise and $k \in \mathbb{N}$ is the time index.

C. Formulation of the Moving Horizon Estimator (MHE)

Let us consider now the generic formulation of the Moving Horizon Estimator (MHE). Assume that the state vector x_k has to be estimated at instant $k \geq N$ using the latest $N+1$ measurements collected within the "sliding horizon" $[k-N,k]$: $y_{k-N}^k = \begin{pmatrix} y_{k-N}^T & y_{k-N+1}^T & \cdots & y_k^T \end{pmatrix}^T$. Denote $\hat{x}_{k-N|k}^-$ an a priori estimate of x_{k-N} and $\hat{x}_{k-N|k}$ an a posteriori estimate, or shortly an estimate, of x_{k-N}, computed before and after the acquisition of measurement y_k respectively.

The MHE in both deterministic and stochastic frameworks can be described by the following generic formulation:

$$\min_{\hat{x}_{k-N|k},\{\hat{w}_{k-N+i|k}\}_{i=0}^{N-1}} \mathbf{J}_k(\hat{x}_{k-N|k},\{\hat{w}_{k-N+i|k}\}_{i=0}^{N-1},\hat{x}_{k-N|k}^-,y_{k-N}^k) \quad (2a)$$

$$\mathbf{J}_k = \left\|\hat{x}_{k-N|k}-\hat{x}_{k-N|k}^-\right\|_{\mathbf{P}}^2 + \sum_{i=0}^{N-1}\left\|\hat{w}_{k-N+i|k}\right\|_{\mathbf{Q}}^2$$

$$+ \sum_{i=0}^{N}\left\|y_{k-N+i}-h(\hat{x}_{k-N+i|k})\right\|_{\mathbf{R}}^2 \quad (2b)$$

$$\text{s. t.} \quad \hat{x}_{k-N+i+1|k} = f(\hat{x}_{k-N+i|k}) + \hat{w}_{k-N+i|k}, \ \forall i \in [0,N-1] \quad (2c)$$

$$\hat{x}_{k-N|k} \in \mathbb{X}, \ \hat{w}_{k-N+i|k} \in \mathbb{W}, \ \forall i \in [0,N-1] \quad (2d)$$

Denote $(\hat{x}_{k-N|k}^\circ,\{\hat{w}_{k-N+i|k}^\circ\}_{i=0}^{N-1})$ the optimal solution of the problem (2) computed at instant k. The a priori estimate at the beginning of the horizon $\hat{x}_{k-N|k}^-$ is computed using $(\hat{x}_{k-N-1|k-1}^\circ,\hat{w}_{k-N-1|k-1}^\circ)$ given by the MHE at the previous instant $k-1$ and (2c). $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{R}$ are positive definite weight matrices. The estimate of x_k computed at instant k given by the MHE is denoted by $\hat{x}_{k|k}$ and is computed using $(\hat{x}_{k-N|k}^\circ,\{\hat{w}_{k-N+i|k}^\circ\}_{i=0}^{N-1})$ propagated through the state equation (2c). For $k < N$, the formulation is similar to (2) but with the increasing horizon equal to k at each instant k instead.

III. DIFFERENCES BETWEEN MHE IN DETERMINISTIC AND STOCHASTIC FRAMEWORKS

A. Assumptions

As previously stated, the main difference between MHE formulations in both frameworks lies in the representation of uncertainties and the underlying assumptions. In the deterministic framework, the initial state x_{k-N} at the begining of the horizon, the process noise w_k and the measurement noise v_k are considered as deterministic variables of unknown characteristics taking their values in known compact sets [5][6]. While in the stochastic framework as in [2][7][8], these parameters are considered as random variables and assumptions on their statistics are made. These two different main assumptions also impact the way other design parameters are tuned in the MHE algorithm.

B. Parameter tuning

The main parameters of the MHE algorithm are the weight matrices $\mathbf{P},\mathbf{Q},\mathbf{R}$ used in the cost function (2b).

In the stochastic framework, these matrices are usually defined as $\mathbf{Q} = Q^{-1}$ and $\mathbf{R} = R^{-1}$, where Q and R are the process noise and the measurement noise covariance matrices respectively.

The matrix $\mathbf{P}$ is defined as the inverse of the a priori error covariance matrix $P_{k-N|k}^{-1}$ which should account for information neglected before the horizon of MHE. Therefore $P_{k-N|k}^{-1}$ is usually updated using the Extended Kalman Filter equations [7] or sampling-based filters [8].

In the deterministic framework, the work of [5][6] proposes to define the weight matrix $\mathbf{P}$ as a constant matrix such as $\mathbf{P} = p\mathbf{I}_{n_x}$ where $p \in \mathbb{R}^{+*}$ and $\mathbf{I}_{n_x}$ is the identity matrix of dimension n_x. The choice of the value for this parameter p is constrained by stability conditions, as it will be discussed in the next section.

C. Stability of the Dynamics of the Estimation Errors

Let us introduce first the definitions concerning the stability of the dynamics of the estimation errors of an estimator.

Definition 1: (Robust Global Asymptotic Stability) Denote $\hat{x}_0^-$ an a priori initial estimate. The dynamics of the estimation error $(\hat{x}_k - x_k)$ for the system (1) $\forall k \in \mathbb{N}$ is Robustly Globally Asymptotically Stable (RGAS) iff $\forall x_0 \in \mathbb{X}$, $\forall \hat{x}_0^- \in \mathbb{X}$, $\forall w_0^{k-1} \in \mathbb{W}^k$ and $\forall v_0^k \in \mathbb{V}^{k+1}$, there exist a class-KL function α and class-K functions δ_w and δ_v such that $\forall k \in \mathbb{N}$ $\|\hat{x}_k - x_k\|^2 \leq \alpha(\|\hat{x}_0^- - x_0\|^2,k) + \delta_w(\|w_0^{k-1}\|^2) + \delta_v(\|v_0^k\|^2)$.

Definition 2: (Observability in $N+1$ steps) Define the observation map of the system (1) on a window of length $N+1$ by

$$F(x_{k-N},w_{k-N}^{k-1}) \triangleq \begin{pmatrix} h(x_{k-N}) \\ h(f^{w_{k-N}}(x_{k-N})) \\ \vdots \\ h \circ f^{w_{k-1}} \circ \cdots \circ f^{w_{k-N}}(x_{k-N}) \end{pmatrix} \quad (3)$$

where $\circ$ denotes function composition and $f^{w_i}(x_i) \triangleq f(x_i) + w_i = x_{i+1}$. Denote also

$$f^{w_{k-N}^{k-N+i}}(x_i) = f^{w_{k-N+i}} \circ \cdots \circ f^{w_{k-N}}(x_{k-N}), \ i \in [0,N-1] \quad (4)$$

The system (1) is said to be observable in $N+1$ steps if this condition is satisfied: $\forall(x',x'') \in \mathbb{X}^2$, there exists a K-function $\phi(\cdot)$ such that

$$\phi(\|x' - x''\|^2) \leq \|F(x',\mathbf{0}_{n_xN\times 1}) - F(x'',\mathbf{0}_{n_xN\times 1})\|^2 \quad (5)$$

Let us consider now the work in the literature concerning the stability of the dynamics of the estimation errors for nonlinear MHE in both frameworks.

1) In Stochastic Framework: It is proven in [7] that the dynamics of the estimation errors of the MHE in this framework is RGAS if the system (1) is observable in $N+1$ steps and the term $\left\|\hat{x}_{k-N|k} - \hat{x}_{k-N|k}^-\right\|_{\mathbf{P}}^2$ which is called an "approximated" arrival cost in [7] is lower than the "exact" arrival cost which is equal to the sum of the cost function concerning every measurements from $k = 0$ to $k-N-1$, i.e. y_0^{k-N-1}. As stated in [5], this condition is difficult to be used in practice since the exact arrival cost must be computed at each instant which leads to large computational effort. Despite this difficulty to cope with this stability condition in practice, MHE formulation in the stochastic framework has been widely used ans studied in many works [1][8][9][10].

2) In Deterministic Framework: It is shown in [5] that the dynamics of the estimation errors of the MHE in the deterministic framework is RGAS if the weight matrix **P** is adequately chosen. An adequate **P** can be calculated if the system (1) satisfies the following assumptions:

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

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

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

(A4) The system (1) is observable in $N+1$ steps

(A5) The system (1) has finite sensitivity, i.e. the minimum of the class-K function $\phi(\cdot)$ in (5) exists. Denote δ the sensitivity parameter of the system (1) defined by

$$\delta = \inf_{(x', x'') \in \mathbb{X}^2, x' \neq x''} \frac{\phi(\|x' - x''\|^2)}{\|x' - x''\|^2} > 0 \qquad (6)$$

Assumption (A5) ensures that a variation of the state vector at the beginning of the horizon x_{k-N} induces a variation in the observation map $F(x_{k-N}, w_{k-N}^{k-1})$ which corresponds to the measurement sequence y_{k-N}^k in case of zero measurement noise. The larger δ is, the more observable the system is [6] (the smaller the number of the measurements needed for building the estimate).

The theorem on the convergence of the estimation errors of the MHE is proposed in [5] for $\mathbf{P} = p\mathbf{I}_{n_x}$ with $p \in \mathbb{R}^{+*}$ as follows:

Theorem 1: For a discrete-time nonlinear system under bounded additive noises as in (1) satisfying assumptions (A1)-(A5), the dynamics of the estimation errors of the MHE designed in the deterministic framework defined in (2) is RGAS if the design parameter p is chosen such that

$$c_1 p + \frac{(L_f^x)^2 c_2 p}{p + c_3 \delta} < 1 \qquad (7)$$

where c_1, c_2 and c_3 are suitable positive constants.

This theorem provides a stability condition that is tractable for practical use of MHE in the deterministic framework. However it would require the explicit expression of the constants c_1, c_2 and c_3 which were not provided in the literature.

These expressions are given here for the case of diagonalizable matrices **Q** and **R** (see details in appendix A).

$$c_1 = 6/\lambda_{\mathbf{Q}}^{min}, \; c_2 = 12, \; c_3 = 2\lambda_{\mathbf{R}}^{min}/3 \qquad (8)$$

IV. ROBUSTNESS ASSESSMENT OF MHE AGAINST MODEL ERRORS: CASE STUDY

As a case study, we consider a problem of space debris trajectory tracking during atmospheric re-entry on Earth from measurements provided by a radar ground station. For the sake of simplicity, a simplified one dimensional problem is considered as presented in figure 1. In this problem, the
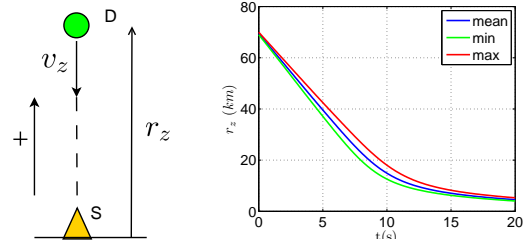

Fig. 1: Space Debris Trajectory in a 1 dimensional case, D: debris and S: radar station (left). Evolutions of the position of the simulated trajectories at each instant in minimum, mean and maximum (right).

acceleration of the debris is governed by gravitational and aerodynamic drag forces. This latter depends on a parameter $\beta(t)$ (ballistic coefficient) which is unknown and has high variation with respect to time, which makes it difficult to model. This estimation problem is therefore subject to high model errors which can even lead to filter divergence.

A. Simulation of Real Trajectories and Measurements

Monte Carlo simulations of 100 trajectories of the debris during 20 s starting from an altitude close to 70 km have been performed assuming a spherical debris, with a varying $\beta(t)$, as done in [3]. The minimum, the mean and the maximum of the altitude of the debris at each instant among all the simulated trajectories are shown in figure 1. Radar measurements associated to these trajectories have been computed. In this simplified one dimensional problem, they consist in the distances between the debris and the radar ground station. The evolution of the ballistic coefficient $\beta(t)$ of the debris is supposed to be unknown to the estimators, which would be the case in reality. This allows to assess the robustness of the estimators with respect to model errors.

The performance of the following estimators will be studied: the Extended Kalman Filter (EKF), the Unscented Kalman Filter (UKF), the Regularized Particle Filter (RPF) and the Moving Horizon Estimators (MHE) in the deterministic framework and in the stochastic framework.

B. Estimation Model

Define as $x_k = \begin{pmatrix} r_{z_k} & v_{z_k} & a_{z_k} \end{pmatrix}^T$ the state at instant k that has to be estimated. To derive an estimation model for the estimators, let us assume that β is constant over a sampling period T_s. The following state equation is obtained [11]:

$$x_{k+1} = f(x_k) + w_k = \begin{pmatrix} 1 & T_s & \frac{T_s^2}{2} \\ 0 & 1 & T_s \\ 0 & 0 & 1 \end{pmatrix} x_k + \begin{pmatrix} \frac{T_s^3}{6} \\ \frac{T_s^2}{2} \\ T_s \end{pmatrix} f_{z_k} + w_k \quad (9)$$

where $f_{z_k} \triangleq \dot{a}_z(t_k)|_{\beta=cst} = (-c_2 v_{z_k} + \frac{2a_{z_k}}{v_{z_k}})(a_{z_k} + g_E)$, with g_E the Earth gravity constant, and w_k is a modeled as a discrete-time zero-mean truncated gaussian white noise of covariance

Estimator	$ARMSE(\hat{r}_z)$ (m)	
	$\sqrt{\tilde{q}_z}T_s = 15$	$\sqrt{\tilde{q}_z}T_s = 1.5$
EKF	5.90	7.73
UKF	5.91	7.70
RPF	7.81	×
MHE deter.	8.22	8.31
MHE sto.	6.45	6.38

TABLE I: ARMSE for each estimator for the two values of $\tilde{q}_z$

Fig. 2: RMSE for each estimator for the two values of $\tilde{q}_z$

matrix Q. This matrix can be parameterized as $Q = \mathcal{Q}(T_s).\tilde{q}_z$ where the matrix $\mathcal{Q}(T_s)$ can be found in [11] and the scalar $\tilde{q}_z > 0$ should account for both discretization error and model errors, from the fact that β is not constant over a sampling period.

The measurement equation is simply $y_k = r_{z_k} + v_k$ where the measurement noise v_k is modeled as a zero-mean truncated Gaussian white noise of covariance $R = 10^2$ m^2.

C. Tuning of the Estimators

For each run, the following parameters are given to each estimator. The sampling period of the estimator $T_s = 0.1$ s is chosen which is equal to the period of radar measurements. From simulated trajectories, we know that $\tilde{q}_z$ should be chosen such that $\sqrt{\tilde{q}_z}T_s = 15$ m/s^2 to capture the variation of β during the sampling period for the first 10 seconds of the trajectories and that $\tilde{q}_z$ should be chosen such that $\sqrt{\tilde{q}_z}T_s = 1.5$ m/s^2 to capture the variation of β over a sampling period during the last 10 seconds of the trajectories. To study the robustness of the estimators against bad process noise tuning, $\tilde{q}_z$ will be considered constant over the whole trajectory and these two values will be tested for the estimators to analyze their performance. Since for the MHEs in both frameworks constraints can be imposed during the optimization, it is supposed that their estimate $\hat{x} \in \mathbb{X}$ where $\mathbb{X} \triangleq ([0,100]$ $km \times [-10000,10]$ $m/s \times [-20,2000]$ $m/s^2)$. Concerning the process noise estimate, to have less number of optimization parameters, it is supposed that $\hat{w}_k = \begin{pmatrix} \frac{T_s^2}{2} & T_s & 1 \end{pmatrix}^T \hat{w}_{a_z,k}$ where $\hat{w}_{a_z,k} \in [-40,40]$ m/s^2 is the only process noise parameter to be estimated at the instant k. The horizon length $N = 6$ is chosen. For the MHE in the deterministic framework, $p = 5 \cdot 10^{-9}$ is chosen, allowing (7) to be verified, and stability to be guaranted. More details on the calculation of this value for p can be found in [11], as well as on tuning parameters for the UKF and RPF used as comparison.

D. Performance Analysis

Let us first define the Root Mean Square Error (RMSE) and the Average Root Mean Square Error (ARMSE) of the ith component of the estimate $\hat{x}_{i,k}$

$$RMSE(\hat{x}_{i,k}) = \sqrt{\sum_{n \in \mathbb{M}} \frac{(\hat{x}_{i,k}^n - x_{i,k}^n)^2}{N_{\mathbb{M}}}}, \quad ARMSE(\hat{x}_{i,k}) = \sum_{k=0}^{k_f} \frac{RMSE(\hat{x}_{i,k})}{k_f + 1} \quad (10)$$

where $\mathbb{M}$ is the set of indexes of *the non-divergent runs*, $N_{\mathbb{M}}$ the number of its elements and k_f is the discrete-time index of the final instant t_f. The ARMSE is defined as the RMSE
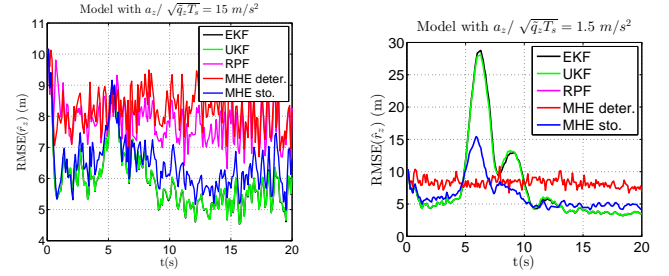
average over all time instants. The RMSE of the position estimate $\hat{r}_z$ for both process noise cases are presented in figures 2. The ARMSE of $\hat{r}_z$ are presented in Table I.

E. Result Analysis

We observe that when the process noise is adequately chosen, the EKF and the UKF provide the smallest estimation errors. However, when the dynamics of the system is not well known, they may provide high errors and therefore should not be used. The RPF is very sensitive to process noise value and did not converge when high model error are considered.

On the other hand, the MHEs are more robust against model errors induced from a bad choice of the process noise parameter. We observe also that the MHE in the stochastic framework provides a peak of errors when the process noise parameter is not well chosen, as for other stochastic estimators. This is not the case for the MHE in the deterministic framework. Even though it provides an overall estimation error higher than the MHE in the stochastic framework, the MHE in the deterministic framework is the only estimator that provides the same level of accuracy for any process noise parameter choice and is therefore the most robust for this tracking problem.

V. CONCLUSIONS AND PERSPECTIVE

In this paper, a discussion has been proposed on the categorization of the formulations of the nonlinear MHE of the literature. Two frameworks have been considered, stochastic and deterministic, depending on the assumptions used to model uncertainties relative to noises and optimization parameters.

The MHE in the stochastic framework proposed to use a weight matrix on the a priori initial state at the beginning of the horizon calculated and propagated using EKF equations. This choice is easy to implement, and has been shown in many works to be efficient for state estimation but conditions to guarantee convergence of the estimates may be difficult to cope with. In the MHE algorithm considered in the deterministic framework, this weight matrix can be more easily defined to satisfy a stability condition. To make it directly usable in practice, explicit formulation of all the terms involved in this condition have been provided in this paper, completing the stability theorem proposed in the literature.

This paper also proposed an assessment of the robustness of these estimators with respect to model errors, through a simulation example, showing the great robustness of both MHEs, and especially the deterministic one.

APPENDIX

We would like to have expressions of the constants c_1, c_2 and c_3 for the convergence of the estimation errors of the MHE defined in (2) when the weight matrix on the arrival cost $\mathbf{P} = p\mathbf{I}_{n_x}$ and $\mathbf{Q}$ and $\mathbf{R}$ are symmetric positive definite diagonalizable matrices. Following [6], we calculate an upper bound and a lower bound on the optimal cost $\mathbf{J}_k^\circ$ defined as the cost corresponding to the optimal estimate $\hat{x}_{k-N|k}^\circ$, i.e. $\mathbf{J}_k^\circ \triangleq \mathbf{J}(\hat{x}_{k-N|k}^\circ, \hat{x}_{k-N|k}^-, y_k^{k-N})$. Then, combining the upper bound to the lower bound gives the dynamics of the estimation errors and of the estimated process noise sequence.

Using the definition of the observation map in (3), the cost function becomes

$$\mathbf{J}_k = p\left\|\hat{x}_{k-N|k} - \hat{x}_{k-N|k}^-\right\|^2 + \sum_{i=0}^{N-1}\left\|\hat{w}_{k-N+i|k}\right\|_{\mathbf{Q}}^2$$
$$+ \left\|y_{k-N}^k - F(\hat{x}_{k-N|k}, \hat{w}_{k-N|k}^{k-1})\right\|_{\mathbf{R}}^2 \quad (11)$$

Let us define as in [6] $\forall x'_{k-N} \in \mathbb{X}$, $\forall w_{k-N+i} \in \mathbb{W}$, $i = 0,1,\dots,N-1$

$$D_w(x'_{k-N}) \triangleq \begin{pmatrix} \frac{\partial y_{k-N}}{\partial w_{k-N}} & \frac{\partial y_{k-N}}{\partial w_{k-N+1}} & \cdots & \frac{\partial y_{k-N}}{\partial w_{k-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_k}{\partial w_{k-N}} & \frac{\partial y_k}{\partial w_{k-N+1}} & \cdots & \frac{\partial y_k}{\partial w_{k-1}} \end{pmatrix} \quad (12)$$

Using (4), we have

$$D_w(x'_{k-N}) = \begin{pmatrix} 0 & \cdots & 0 \\ \frac{\partial(h \circ f_{(1)}^{w_{k-N}})}{\partial w_{k-N}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ \frac{\partial(h \circ f_{(N)}^{w_{k-N}^{k-1}})}{\partial w_{k-N}} & \cdots & \frac{\partial(h \circ f_{(N)}^{w_{k-N}^{k-1}})}{\partial w_{k-1}} \end{pmatrix} \quad (13)$$

$$\Delta_w \triangleq \max_{x'_{k-N} \in \mathbb{X}} \|D_w(x'_{k-N})\| \,|_{w_{k-N}^{k-1} = \mathbf{0}_{n_x N \times 1}} \quad (14)$$

Define $\bar{k}$ such that $\bar{k} \triangleq \max_{i=1,\dots,N} \bar{k}_i$ where $\bar{k}_i > 0$ are suitable scalars such that $\bar{k}_i\left\|w_{k-N}^{k-N+i-1}\right\| \geq$

$$\left\|\frac{\partial(h \circ f_{(N)}^{w_{k-N}^{k-N+i-1}})}{\partial w_{k-N}^{k-N+i-1}}\big|_{w_{k-N}^{k-N+i-1}} - \frac{\partial(h \circ f_{(N)}^{w_{k-N}^{k-N+i-1}})}{\partial w_{k-N}^{k-N+i-1}}\big|_{\mathbf{0}_{n_x N \times 1}}\right\| \quad (15)$$

The constant $\bar{k}_i$ exists as we are dealing with compositions of $\mathscr{C}^2$ functions. Although it may not be easily computed, its value only affects the asymptotic bound on the estimation error but not its stability.

Upper Bound on the Optimal Cost $\mathbf{J}_k^\circ$

Since $\mathbf{J}_k^\circ$ is the optimal solution of (2), it is smaller than the cost calculated at the real state x_{k-N} with zero process noise $\mathbf{0}_{n_x N \times 1}$. In other words,

$$\mathbf{J}_k^\circ \leq p\left\|x_{k-N} - \hat{x}_{k-N|k}^-\right\|^2 + \left\|y_k^{k-N} - F(x_{k-N}, \mathbf{0}_{n_x N \times 1})\right\|_{\mathbf{R}}^2$$
$$\leq p\left\|x_{k-N} - \hat{x}_{k-N|k}^-\right\|^2 + \lambda_{\mathbf{R}}^{max}\left\|y_k^{k-N} - F(x_{k-N}, \mathbf{0}_{n_x N \times 1})\right\|^2$$

According to [6], the last term on the r.h.s of the above equation follows

$$\left\|y_k^{k-N} - F(x_{k-N}, \mathbf{0}_{n_x N \times 1})\right\|^2 \leq c_{Ale,N}^2 \quad (16)$$

where

$$c_{Ale,N} \triangleq \Delta_w\sqrt{N}r_w + \sqrt{N+1}r_v + \frac{\bar{k}}{2}\sqrt{\frac{N(N+1)(2N+1)}{6}}r_w^2 \quad (17)$$

Therefore, we have an upper bound on the optimal cost $\mathbf{J}_k^\circ$ as follows

$$J_k^\circ \leq p\left\|x_{k-N} - \hat{x}_{k-N|k-1}\right\|^2 + \lambda_{\mathbf{R}}^{max}c_{Ale,N}^2 \quad (18)$$

Lower Bound on the Optimal Cost $\mathbf{J}_k^\circ$

Denote, for $i = 0,\dots,N$, $\hat{x}_{k-N}$ any estimate of x_{k-N} and $\hat{w}_{k-N}^{k-1}$ any sequence of estimated process noises from $k-N$ to $k-1$. We would like to compute the cost in (11) for any solution $(\hat{x}_{k-N}, \hat{w}_{k-N}^{k-1})$, not necessarily computed at k. Using the triangle inequality proposed in [6], one obtains

$$\left\|F(x_{k-N}, \mathbf{0}_{n_x N \times 1}) - F(\hat{x}_{k-N}, \mathbf{0}_{n_x N \times 1})\right\|^2$$
$$\leq 3\left\|F(x_{k-N}, \mathbf{0}_{n_x N \times 1}) - y_{k-N}^k\right\|^2 + 3\left\|y_{k-N}^k - F(\hat{x}_{k-N}, \hat{w}_{k-N}^{k-1})\right\|^2$$
$$+ 3\left\|F(\hat{x}_{k-N}, \hat{w}_{k-N}^{k-1}) - F(\hat{x}_{k-N}, \mathbf{0}_{n_x N \times 1})\right\|^2$$

Using (5) and (16) and the fact that $\mathbf{R}$ is diagonalizable, the third term of the cost function (11) becomes

$$\left\|y_k^{k-N} - F(x_{k-N}, \mathbf{0}_{n_x N \times 1})\right\|_{\mathbf{R}}^2 \geq \lambda_{\mathbf{R}}^{min}(\frac{1}{3}\phi(\|x_{k-N} - \hat{x}_{k-N}\|^2)$$
$$- c_{Ale,N}^2 - \left\|F(\hat{x}_{k-N}, \hat{w}_{k-N}^{k-1}) - F(\hat{x}_{k-N}, \mathbf{0}_{n_x N \times 1})\right\|^2) \quad (19)$$

According to [6], we have

$$\left\|F(\hat{x}_{k-N}, \hat{w}_{k-N}^{k-1}) - F(\hat{x}_{k-N}, \mathbf{0}_{n_x N \times 1})\right\|^2$$
$$\leq (\Delta_w\sqrt{N}r_w + \frac{\bar{k}}{2}\sqrt{\frac{N(N+1)(2N+1)}{6}}r_w^2)^2$$

Hence, (19) becomes

$$\left\|y_k^{k-N} - F(x_{k-N}, \mathbf{0}_{n_x N \times 1})\right\|_{\mathbf{R}}^2 \geq \lambda_{\mathbf{R}}^{min}[\frac{1}{3}\phi(\|x_{k-N} - \hat{x}_{k-N}\|^2)$$
$$- c_{Ale,N}^2 - (\Delta_w\sqrt{N}r_w + \frac{\bar{k}}{2}\sqrt{\frac{N(N+1)(2N+1)}{6}}r_w^2)^2] \quad (20)$$

Using the triangle inequality, we obtain $p\left\|\hat{x}_{k-N} - \hat{x}_{k-N|k-1}\right\|^2 \geq$

$$\frac{1}{2}p\|x_{k-N} - \hat{x}_{k-N}\|^2 - p\left\|x_{k-N} - \hat{x}_{k-N|k-1}\right\|^2 \quad (21)$$

Again, since $\mathbf{Q}$ is diagonalizable we can write

$$\sum_{i=0}^{N-1}\left\|\hat{w}_{k-N+i|k}\right\|_{\mathbf{Q}}^2 \geq \lambda_{\mathbf{Q}}^{min}\sum_{i=0}^{N-1}\|\hat{w}_{k-N+i}\|^2 \quad (22)$$

Define by $e_{k-N|k}^\circ \triangleq x_{k-N} - \hat{x}_{k-N|k}^\circ$ the estimation error associated to the optimal estimate of x_{k-N} given by the MHE computed at instant k and the associated optimal process

noise sequence $\{\hat{w}^{\circ}{}_{k-N|k}^{k-1}\}$. Remark that $\sum_{i=0}^{N-1}\left\|\hat{w}^{\circ}{}_{k-N+i|k}^{k-1}\right\|^2 = \left\|\hat{w}^{\circ}{}_{k-N|k}^{k-1}\right\|^2$. Combine (20)-(22) to have the lower bound:

$$\mathbf{J}_k^{\circ} \geq \frac{1}{2}p\left\|e^{\circ}_{k-N|k}\right\|^2 - p\left\|x_{k-N}-\hat{x}^{-}_{k-N|k}\right\|^2 \tag{23}$$
$$+ \frac{1}{3}\lambda_{\mathbf{R}}^{min}\phi(\left\|e^{\circ}_{k-N|k}\right\|^2) - \lambda_{\mathbf{R}}^{min}c'^2 + \lambda_{\mathbf{Q}}^{min}\left\|\hat{w}^{\circ}{}_{k-N|k}^{k-1}\right\|^2$$

where

$$c'^2 = c_{Ale,N}^2 + (\Delta_w\sqrt{N}r_w + \frac{\bar{k}}{2}\sqrt{\frac{N(N+1)(2N+1)}{6}}r_w^2)^2 \tag{24}$$

Now, let us combine the upper and the lower bounds on the optimal cost to have the dynamics of the estimation errors e°_{k-N} and the associated estimated process noise $\hat{w}^{\circ}{}_{k-N}^{k-1}$.

Dynamics of the Estimation Errors of the MHE

Combining the upper bound in (18) to the lower bound in (23), and using (6), one gets:

$$\left(\frac{p}{2}+\frac{\lambda_{\mathbf{R}}^{min}}{3}\delta\right)\left\|e^{\circ}_{k-N|k}\right\|^2 + \lambda_{\mathbf{Q}}^{min}\left\|\hat{w}^{\circ}{}_{k-N|k}^{k-1}\right\|^2$$
$$\leq 2p\left\|x_{k-N}-\hat{x}_{k-N|k-1}\right\|^2 + \mathbf{c}^2 \tag{25}$$

where $\quad \mathbf{c}^2 = \lambda_{\mathbf{R}}^{max}c_{Ale,N}^2 + \lambda_{\mathbf{R}}^{min}c'^2 \tag{26}$

Now, using the Lipschitz continuity of f, we have $\left\|x_{k-N}-\hat{x}^{-}_{k-N|k}\right\|^2$

$$= \left\|f(x_{k-N-1})-f(\hat{x}^{\circ}_{k-N-1|k-1})-\hat{w}^{\circ}_{k-N-1|k-1}+w_{k-N-1}\right\|^2$$
$$\leq 3(L_f^x)^2\left\|e^{\circ}_{k-N-1|k-1}\right\|^2 + 3\left\|\hat{w}^{\circ}_{k-N-1|k-1}\right\|^2 + 3r_w^2 \tag{27}$$

Replace (27) into (25) to get

$$\left(\frac{p}{2}+\frac{\lambda_{\mathbf{R}}^{min}}{3}\delta\right)\left\|e^{\circ}_{k-N|k}\right\|^2 + \lambda_{\mathbf{Q}}^{min}\sum_{i=0}^{N-1}\left\|\hat{w}^{\circ}_{k-N+i|k}\right\|^2 \leq$$
$$2p\left(3(L_f^x)^2\left\|e^{\circ}_{k-N-1|k-1}\right\|^2 + 3\left\|\hat{w}^{\circ}_{k-N-1|k-1}\right\|^2 + 3r_w^2\right) + \mathbf{c}^2$$

Using $\left\|\hat{w}^{\circ}_{k-N-1|k-1}\right\|^2 \leq \left\|\hat{w}^{\circ}{}_{k-N-1|k-1}^{k-2}\right\|^2$, we obtain

$$\left\|e^{\circ}_{k-N|k}\right\|^2 \leq \frac{1}{p+\frac{2\lambda_{\mathbf{R}}^{min}}{3}\delta}(12p(L_f^x)^2\left\|e^{\circ}_{k-N-1|k-1}\right\|^2$$
$$+12p\left\|\hat{w}^{\circ}{}_{k-N-1|k-1}^{k-2}\right\|^2 + 12pr_w^2 + 2\mathbf{c}^2)$$
$$\left\|\hat{w}^{\circ}{}_{k-N|k}^{k-1}\right\|^2 \leq \frac{1}{\lambda_{\mathbf{Q}}^{min}}(6p(L_f^x)^2\left\|e^{\circ}_{k-N-1|k-1}\right\|^2$$
$$+6p\left\|\hat{w}^{\circ}{}_{k-N-1|k-1}^{k-2}\right\|^2 + 6pr_w^2 + \mathbf{c}^2)$$

In other words, the stability theorem 1 proposed in [5] is obtained with $\left\|e^{\circ}_{k-N|k}\right\|^2 \leq \xi_{k-N}$ and $\left\|\hat{w}^{\circ}{}_{k-N|k}^{k-1}\right\|^2 \leq \omega_{k-N}$ where the sequences $\{\xi_k\}$ and $\{\omega_k\}$ are generated by

the linear system $\begin{pmatrix}\omega_{k+1} & \xi_{k+1}\end{pmatrix}^T = A(p,\delta).\begin{pmatrix}\omega_k & \xi_k\end{pmatrix}^T + B(p,\delta).\begin{pmatrix}r_w^2 & r_v^2\end{pmatrix}^T$ with

$$A(p,\delta) \triangleq \begin{pmatrix} c_1p & c_1(L_f^x)^2p \\ c_2p/(p+c_3\delta) & c_2(L_f^x)^2p/(p+c_3\delta) \end{pmatrix}$$
$$B(p,\delta) \triangleq \begin{pmatrix} c_4p+c_5 & c_6 \\ (c_7p+c_8)/(p+c_3\delta) & c_9/(p+c_3\delta) \end{pmatrix}$$

where thanks to the calculations in this section, we find $c_1 = \frac{6}{\lambda_{\mathbf{Q}}^{min}}$, $c_2 = 12$, $c_3 = \frac{2}{3}\lambda_{\mathbf{R}}^{min}$. The expressions of $c_4,\ldots,c_9$ can be computed using the expression of $\mathbf{c}^2$. To compute $\mathbf{c}$, one has to compute the values of Δ_w in (14) and $\bar{k}_i$ in (15). One way to compute Δ_w when an analytical expression of D_w is available consists in making hypotheses on the ranges of possible values of the state and the process noise. To compute $\bar{k}_i$, remark that $\bar{k}_i$ is in fact a Lipschitz constant of the function $\partial(h\circ f_i^{w_{k-N}^{k-i}})/\partial w_{k-N}^{k-1}$. It can be computed using an analytical expression of the Jacobian matrix of $\partial(h\circ f_i^{w_{k-N}^{k-i}})/\partial w_{k-N}^{k-1}$ and then making hypotheses on the ranges of possible values of the state and the process noise to compute its norm, i.e. the Lipschitz constant.

REFERENCES

[1] 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, pp. 2451–2460, 2005.

[2] J. B. Rawlings and B. R. Bakshi, "Particle filtering and moving horizon estimation," *Computers & chemical engineering*, vol. 30, no. 10, pp. 1529–1541, 2006.

[3] R. Suwantong, S. Bertrand, D. Dumur, and D. Beauvois, "Moving Horizon Estimation with Pre-Estimation (MHE-PE) for 3D Space Debris Tracking during Atmospheric Re-entries," in *IEEE Conference on Decision and Control*, 2014.

[4] A. Alessandri, M. Baglietto, and G. Battistelli, "Robust receding-horizon estimation for discrete-time linear systems in the presence of bounded uncertainties," in *IEEE Conference on Decision and Control*, 2005.

[5] A. Alessandri, M. Baglietto, G. Battistelli, and V. Zavala, "Advances in moving horizon estimation for nonlinear systems," in *IEEE Conference on Decision and Control*, 2010.

[6] 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, no. 7, pp. 1753–1765, 2008.

[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, pp. 246–258, 2003.

[8] S. Ungarala, "Computing arrival cost parameters in moving horizon estimation using sampling based filters," *Journal of Process Control*, vol. 19, no. 9, pp. 1576–1588, 2009.

[9] R. López-Negrete, S. C. Patwardhan, and L. T. Biegler, "Constrained particle filter approach to approximate the arrival cost in moving horizon estimation," *Journal of Process Control*, vol. 21, no. 6, pp. 909–919, 2011.

[10] C. C. Qu and J. Hahn, "Computation of arrival cost for moving horizon estimation via unscented Kalman filtering," *Journal of Process Control*, vol. 19, no. 2, pp. 358–363, 2009.

[11] R. Suwantong, "New structure for moving horizon estimators. application to space debris tracking during the atmospheric re-entries," Ph.D. dissertation, Supélec, 2014.