

Partial Extended Observability Certification and Optimal Design of Moving Horizon Estimators

Mazen Alamir

Abstract—This paper addresses the observability analysis and the optimal design of observation parameters in the presence of noisy measurements and parametric uncertainties. The notion of almost ϵ -observability is introduced and a systematic procedure to assess its satisfaction for a given system with a priori known measurement noise statistics and parameter discrepancy is sketched. Moreover, the concept of observation-target quantities is introduced in order to analyze the precision with which specific chosen expressions of the state and the parameters can be reconstructed. The overall framework is exposed and validated through an illustrative example.

Index Terms—Nonlinear Systems, Moving-Horizon Estimation, Partial Observability, Observation target, Certification via randomized optimization.

I. INTRODUCTION

IN nowadays data-focused period, there are still many engineering problems that require the design of state observers in the traditional sense [8], namely those which are based on knowledge-driven models. Indeed, data-driven models only involve those variables for which sensors are available on-board. Engineering problems however, involve quite often formulations that refer to real-life variables that are not directly accessible via sensors. Knowledge-based models involve parameters with rather badly known values. The lack of precise knowledge of these values added to the measurement noise may have drastic consequences on the quality of the reconstruction process. When facing such situations, the available solutions can be split into the following three categories: In the first category, only measurement noise with zero mean is considered. The observability is then analyzed in noise-free context using Lie-algebra local tools [13] or thanks to structural dedicated high gain [7] or sliding modes [10] observability assumptions. In the second category, parametric uncertainties are considered. Typical frameworks involve interval analysis [12], [17] which is conducted to derive worst-case bounds on the estimation error [2] at the price of structural assumptions that generally involve the existence of bounding behavior for the uncertain system [16]. The third category is related to moving horizon estimation schemes [5], [18], [19], [21]. This approach corresponds to a very rich literature starting from the early works [18], [20] were the foundations of this optimization-based estimation approach have been laid in

Mazen Alamir is with CNRS/Gipsa-lab, Grenoble-INP, University of Grenoble Alpes, France. (e-mail: mazen.alamir@grenoble-inp.fr)

the deterministic setting. Since, many improving refinements have been proposed (see [21] and the reference therein).

As far as uncertainty handling is concerned, one of the more advanced MHE frameworks has been recently proposed in [19] which improved a family of early works [4], [5], [14], [22] that mainly derive a gain-bounded error between the additive bounded disturbance term and the resulting estimation error. The results of [19] extended the previous results to models with non necessarily additive disturbances term via less conservative estimation of the disturbance-to-error gain. The positioning of the present contribution compared to the previously cited MHE works can be summarized by the following items:

- 1) As far as the estimation error bound is concerned in MHE works, existing results are based on a worst-case analysis via Lipschitz-like upper bounds that induce, as acknowledged by the author of [19], too pessimistic (if not useless) results in some cases. *In the present contribution, a different angle of attack is chosen by focusing the interest on the computation of tight probabilistic assessment of the estimation error excursion during the estimation process in the presence of non vanishing uncertainties on the model's parameters. This is done using a formulation that enables probabilistic certification tools to be deployed.*
- 2) In all previous MHE works, the underlying assumption is that the system is globally detectable in the conventional sense. This means that the underlying paradigm is that of total state (or extended state) detectability. The analysis of partial reconstruction of specific functions of the problem's variables and parameters is not a straightforward extension and needs new results to be worked out from scratch. *By opposition, the present contribution is totally built around the concept of partial estimation and the numerical investigation underlines its relevance by showing that different precision levels might be certifiable for different estimation targets $z = T(x, p)$ where T returns an expression that might contain only a part of the state, a part of the parameter vector or some nonlinear combination of x and p such as some designed state-feedback to cite an example. To say it shortly one would be satisfied with the extended state/parameter vector being non precisely reconstructible from the measurement provided that the expression of interest $z = T(x, u)$, called hereafter the observation target, can be reconstructed with acceptable precision.*

The idea of estimating a function of the system's variable and not the whole state is not new as it is in the heart of the investigations regarding the so-called functional observers (see [15], [23], [24] and the references therein). Although the majority of the related works concern linear systems or linear functions of the state to be reconstructed rather than the whole state, some few works apply to nonlinear dynamics. Almost invariably, the solutions amount at constructing a reduced order observer that is build after exhibiting a lower dimensional dynamics having the same dimension as the function to be estimated. The settings are deterministic and the underlying construction heavily relies on structural assumptions on the mathematical equations of the model.

It is worth underlying here that this paper is not about proposing a new numerical scheme to implement MHE. This can be efficiently done by many available tools [6]. The contribution aims at providing a method for the estimation of the size of possible regions of partial indistinguishability. In other words, when the proposed approach fires a high risk of indistinguishability, this risk materializes **regardless of the formulation of any specific observer**. More precisely, the result of the proposed method gives a *probabilistically assessed boundary of the point-wise estimation error should the observer be able to achieve global optimization successfully*.

Finally, it is worth mentioning that, although the proposed framework concerns constant unknown parameters, it obviously also fits the case of slowly varying parameters (within the observation window).

This paper is organized as follows: First some notation and definitions are introduced in Section II together with basic recalls regarding the definition of Moving Horizon Estimators (MHE). Section III introduces the concept of ϵ -observability of an observation target and shows that its satisfaction requires a robust constraints satisfaction condition. This condition is relaxed in Section IV which introduces the less stringent concept of almost ϵ -observability and shows how this concept can be assessed using randomized optimization frameworks. Section V proposes an illustrative example while Section VI concludes the paper and gives some suggestions for future investigation.

II. DEFINITIONS AND NOTATION

We consider nonlinear systems that are governed by a dynamical equation of the form:

$$x_{k+1} = f(x_k, u_k, p) \quad (1)$$

where $x_k \in \mathbb{X} \subset \mathbb{R}^{n_x}$, $u_k \in \mathbb{U} \subset \mathbb{R}^{n_u}$ stand for the values, at sampling instant k , of the state and the input vectors respectively while $p \in \mathbb{P}$ stands for the vector of imperfectly known parameters of the model. It is also assumed that some vector of n_y measurements (including the input) is available that is linked to the state and the control vector through noisy measurement, namely:

$$y_k = h(x_k, u_k, p) + \nu_k \quad (2)$$

where $y_k \in \mathbb{R}^{n_y}$ stands for the sensor output at instant k which is corrupted by the noise realization ν_k . We assume that one is interested in reconstructing the value of some so-called **observation-target** variable:

$$z = T(x, p) \quad (3)$$

where $T : \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_z}$ is some known map. Given any vector signal

$$s = \begin{bmatrix} s^{[1]} \\ \vdots \\ s^{[n_s]} \end{bmatrix} \in \mathbb{R}^{n_s}$$

the forward/backward profiles of s at some instant k over some window of length N (in terms of sampling periods) is denoted as follows :

$$s_k^+ := [s_k \ s_{k+1} \ \dots \ s_{k+N-1}] \in [\mathbb{R}^{n_s}]^N \quad (4)$$

$$s_k^- := [s_{k-N} \ s_{k-N+1} \ \dots \ s_{k-1}] \in [\mathbb{R}^{n_s}]^N \quad (5)$$

which is a condensed expression gathering the profiles of all the components of s that would be denoted individually according to:

$$s_k^{[i]+} := [s_k^{[i]} \ s_{k+1}^{[i]} \ \dots \ s_{k+N-1}^{[i]}] \in \mathbb{R}^N \quad (6)$$

$$s_k^{[i]-} := [s_{k-N}^{[i]} \ s_{k-N+1}^{[i]} \ \dots \ s_{k-1}^{[i]}] \in \mathbb{R}^N \quad (7)$$

Using the above notation of control profile, it is now possible to define the future state prediction at instant $k+i$ given the current state x_k at instant k , a given control profile $\mathbf{u}_k^+ \in \mathbb{U}^N$ and some vector of parameters p by using the following notation:

$$(\forall i \in \{0, \dots, N\}) \quad \hat{x}_{k+i} = X_i(x_k, \mathbf{u}_k^+, p) \quad (8)$$

which simply refers to the simulation based on (1) with x_k as initial state, \mathbf{u}_k^+ as control profile and p as the parameter vector value. The same notation is used to refer to the predicted noise free output, namely:

$$(\forall i \in \{0, \dots, N\}) \quad \hat{y}_{k+i} = Y_i(x_k, \mathbf{u}_k^+, p) \quad (9)$$

Moreover, the resulting predicted output profile is simply denoted by:

$$\mathbf{Y}(x_k, \mathbf{u}_k^+, p) := [Y_0(x_k, \mathbf{u}_k^+, p) \ \dots \ Y_{N-1}(x_k, \mathbf{u}_k^+, p)] \in [\mathbb{R}^{n_y}]^N \quad (10)$$

Given a candidate initial state ξ at instant $k-N$, an input profile \mathbf{u}_k^- that has been applied over the previous time interval $[k-N, k)$ and a candidate value \hat{p} of parameter vector, a predicted trajectory $\mathbf{Y}(\xi, \mathbf{u}_k^-, \hat{p})$ can be obtained and compared to the truly measured one \mathbf{y}_k^- to define the output prediction error profile by:

$$e_k = \mathbf{y}_k^- - \mathbf{Y}(\xi, \mathbf{u}_k^-, \hat{p}) \in [\mathbb{R}^{n_y}]^N \quad (11)$$

and assuming that the input is measured so that \mathbf{u}_k^- can be viewed as a part of \mathbf{y}_k^- , the last relation can be rewritten shortly as follows:

$$e_k := \mathcal{E}(\xi, \mathbf{y}_k^-, \hat{p}) \quad (12)$$

In order to explicitly acknowledge the presence of measurement errors and parametric uncertainties, we follow here the suggestion made in [1] which amounts at introducing a dead-zone when evaluating the output prediction error, namely, given an error profile $e^{[i]}$ on the output component i , we define the uncertainty-aware distance-to-zero d as follows (in which $[r]_+ = \max\{0, r\}$):

$$d(e^{[i]}, \zeta_i) := \left[\left| \frac{1}{N} \sum_{j=1}^N e_j^{[i]} \right| - \zeta_i \right]_+^r \quad (13)$$

where r is an integer that shape the penalty outside the dead-zone. $r = 1$ is used in the numerical investigation section. Based on this definition, the overall output prediction penalty is defined by:

$$d(e, \zeta) := \sum_{i=1}^{n_y} d(e^{[i]}, \zeta_i) \quad (14)$$

Gathering together equations (14) and (12), it is possible to define at each instant k a cost function:

$$J(\xi, \hat{p} \mid \mathbf{y}_k^-, \zeta) := d(\mathcal{E}(\xi, \mathbf{y}_k^-, \hat{p}), \zeta) \quad (15)$$

This cost function might be used in an extended MHE that first solves the optimization problem

$$(\xi^*, p^*) = \arg \min_{(\xi, p) \in \mathbb{X} \times \mathbb{P}} J(\xi, p \mid \mathbf{y}_k^-, \zeta) \quad (16)$$

and then uses the solution to reconstruct the observation target according to

$$\hat{z}_k = T\left(X_N(\xi^*, \mathbf{u}_k^-, p^*), p^*\right) \quad (17)$$

The *raison d'être* of the dead zone is to make the cost function J equal to 0 when taken at the pair (ξ, p) that issued the output measurement $\mathbf{y} = Y(\xi, \mathbf{u}, p) + \boldsymbol{\nu}$. For this to hold, the dead zone size ζ should be taken high enough. This leads to the following concept:

DEFINITION 1: [Dead-Zones Consistency] The vector of dead zones sizes $\zeta \in \mathbb{R}_+^{n_y}$ is said to be **consistent** if for any admissible state $x \in \mathbb{X}$, any admissible control profile $\mathbf{u} \in \mathbb{U}^N$ and any possible realization of the vector of parameters $p \in \mathbb{P}$, the following equality is satisfied for any realization $\boldsymbol{\nu} \in \mathbb{V}^N$ of the noise profile:

$$J(x, p \mid \mathbf{Y}(x, \mathbf{u}, p) + \boldsymbol{\nu}, \zeta) = 0 \quad (18)$$

which simply means that the correct pair (x, p) corresponds to a vanishing computed cost for any noise and any possible realization of the uncertain parameter vector p .

III. THE ϵ -OBSERVABILITY OF AN OBSERVATION TARGET

In the remainder of this contribution, we shall use the following notation:

$$q := \begin{bmatrix} x \\ p \\ \mathbf{u} \\ \boldsymbol{\nu} \end{bmatrix} =: \begin{bmatrix} q_x \\ q_p \\ q_u \\ q_\nu \end{bmatrix} \in \mathbb{Q} \subset \mathbb{X} \times \mathbb{P} \times \mathbb{U}^N \times \mathbb{V}^N \quad (19)$$

where the second equality is to be considered component-wise, namely, $x = q_x$, $p = q_p$ and so on. This enables to rewrite (18) in the following more condensed form:

$$(\forall q \in \mathbb{Q}) \quad J_1(q, \zeta) := J(q_x, q_p \mid \mathbf{Y}(q_x, q_u, q_p) + q_\nu, \zeta) = 0 \quad (20)$$

Note that each element $q \in \mathbb{Q}$ completely defines a simulation scenario with an associated measurement noise profile. The following definition associates elements of \mathbb{Q} that share the same exogenous information, namely the control input profile and the measurement noise realization:

DEFINITION 2 (COMPARABLE PAIRS): We shall say that two elements $q^{(1)}, q^{(2)} \in \mathbb{Q}$ are **comparable** if and only if they share the components \mathbf{u} and $\boldsymbol{\nu}$. This is denoted as follows:

$$q^{(1)} \bowtie q^{(2)} \Leftrightarrow (q_u^{(1)} = q_u^{(2)}) \text{ and } (q_\nu^{(1)} = q_\nu^{(2)}) \quad (21)$$

Such two elements obviously define two simulations that can differ only by the initial state and/or the vector of parameters while the input profile and the measurement noise are the same.

The next definition introduces the indistinguishability relationship on the set \mathbb{Q} :

DEFINITION 3 (INDISTINGUISHABLE PAIRS):

We shall say that an element $q^{(2)}$ is indistinguishable from $q^{(1)} \in \mathbb{Q}$ (with the notation $q^{(1)} \equiv q^{(2)}$) if and only if it induces 0 output prediction error cost when using as measurement the noisy output generated by $q^{(1)}$, namely

$$J(q_x^{(2)}, q_p^{(2)} \mid \mathbf{Y}(q_x^{(1)}, q_u^{(1)}, w_p^{(1)}) + q_\nu^{(1)}, \zeta) = 0 \quad (22)$$

meaning that $(q_x^{(2)}, q_p^{(2)})$ might as well explain the noisy measurement corresponding to the scenario $q^{(1)}$.

In the ideal situation, no comparable but distinct values of q can be indistinguishable. In this case, one obtains perfect estimation of the state AND the parameter vector. This might be unnecessary in the case where only observation-target variable needs to be reconstructed. That is the reason why the following definition is introduced:

DEFINITION 4 (ϵ -OBSERVABILITY):

Given an observation-target $z = T(x, p)$, we shall say that z is ϵ -observable on \mathbb{Q} if and only if

- 1) There exists a dead-zone size vector $\zeta \in \mathbb{R}_+^{n_y}$ that is consistent over \mathbb{Q} in the sense of Definition 1,
- 2) The following implication holds for any pair $(q^{(1)}, q^{(2)}) \in \mathbb{Q}^2$:

$$\left(q^{(1)} \bowtie q^{(2)} \right) \text{ AND } \left(q^{(1)} \equiv q^{(2)} \right) \Rightarrow \|q_z^{(1)} - q_z^{(2)}\| \leq \epsilon \quad (23)$$

where $q_z := T(q_x, q_p)$ is the observation-target variable associated to q . In other words, only pairs with ϵ -distant observable targets can be both comparable and indistinguishable.

The condition (23) can be written equivalently as follows

thanks to (22):

$$\begin{aligned} & \left(q^{(1)} \bowtie q^{(2)} \right) \quad \text{AND} \quad \left(\|q_z^{(1)} - q_z^{(2)}\| > \epsilon \right) \\ & \quad \quad \quad \downarrow \\ & J(q_x^{(2)}, q_p^{(2)} \mid \mathbf{Y}(q_x^{(1)}, q_u^{(1)}, q_p^{(1)}) + q_\nu^{(1)}, \zeta) \neq 0 \end{aligned} \quad (24)$$

This implication can be written in the following more compact form:

$$\begin{aligned} & \forall q \in \mathbb{Q} \quad \forall (\xi, p) \in \bar{\mathbb{Z}}_\epsilon(q) \\ & J(\xi, p \mid \mathbf{Y}(q_x, q_u, q_p) + q_\nu, \zeta) \neq 0 \end{aligned} \quad (25)$$

where $\bar{\mathbb{Z}}_\epsilon(q)$ is the complement of $\mathbb{Z}_\epsilon(q)$ defined by:

$$\mathbb{Z}_\epsilon(q) := \left\{ (\xi, p) \in \mathbb{X} \times \mathbb{P} \text{ s.t. } \|T(\xi, p) - T(q_x, q_p)\| \leq \epsilon \right\}$$

and introducing the notation:

$$w := \begin{bmatrix} q \\ \xi \\ p \end{bmatrix} \in \mathbb{W} := \mathbb{Q} \times \mathbb{X} \times \mathbb{P} \quad (26)$$

$$\bar{\mathbb{W}}(\epsilon) := \left\{ w = \begin{bmatrix} q \\ \xi \\ p \end{bmatrix} \mid (q, \xi, p) \in \mathbb{Q} \times \bar{\mathbb{Z}}_\epsilon(q) \right\} \quad (27)$$

$$J_2(w, \zeta) := J(\xi, p \mid \mathbf{Y}(q_x, q_u, q_p) + q_\nu, \zeta) \quad (28)$$

one gets the following result:

PROPOSITION 1 (FIRST FORMULATION): An observation-target variable z is ϵ -observable on \mathbb{Q} using the deadzone vector ζ if and only if the following conditions hold true:

- 1) The dead-zone sizes vector is consistent in the sense of Definition 1, namely:

$$(\forall q \in \mathbb{Q}) \quad J_1(q, \zeta) = 0 \quad (29)$$

- 2) The ϵ -distinguishability property holds true, namely

$$(\forall w \in \bar{\mathbb{W}}(\epsilon)) \quad J_2(w, \zeta) \neq 0 \quad (30)$$

where $\bar{\mathbb{W}}(\epsilon)$ is defined by (27) while J_1 and J_2 are defined by (20) and (28) respectively.

Note that the condition (29) can also be written using the notation w_q that extracts the first vector q in the vector w [see (26)] so that one can write the condition (29) as follows:

$$(\forall w \in \bar{\mathbb{W}}(\epsilon)) \quad J_1(w_q, \zeta) = 0 \quad (31)$$

This enables to regroup the two conditions (29) and (30) in a single condition that involves the parameter w , namely:

$$\begin{aligned} & (\forall w \in \bar{\mathbb{W}}(\epsilon)) \\ & C(w, \zeta) := \begin{cases} 0 & \text{if } J_1(w_q, \zeta) = 0 \quad \text{and} \quad J_2(w, \zeta) \neq 0 \\ 1 & \text{otherwise} \end{cases} \end{aligned} \quad (32)$$

This enables Proposition 1 to be reformulated in a more compact form that will be more convenient for the formulation of the probabilistic certification step:

PROPOSITION 2 (SECOND FORMULATION): An observation target variable z is ϵ -observable on \mathbb{W} with the dead-zone vector ζ if and only if the following condition holds true:

$$(\forall w \in \bar{\mathbb{W}}(\epsilon)) \quad C(w, \zeta) = 0 \quad (33)$$

For technical reasons, we need to perform a last transformation by observing that the condition (30) that concerns only those values of w that belong to $\bar{\mathbb{W}}(\epsilon)$ can be transformed into a condition on all possible values of $w \in \mathbb{W} := \mathbb{Q} \times \mathbb{X} \times \mathbb{P}$ by writing

$$(\forall w \in \mathbb{W}) \quad g(w, \epsilon, \zeta) = 0$$

where

$$g(w, \epsilon, \zeta) := \begin{cases} 0 & \text{if } J_1(w_q, \zeta) = 0 \text{ and } w \notin \bar{\mathbb{W}}(\epsilon) \\ 0 & \text{if } J_1(w_q, \zeta) = 0 \text{ and } (w \in \bar{\mathbb{W}}(\epsilon) \text{ and } J_2(w, \zeta) \neq 0) \\ 1 & \text{otherwise} \end{cases} \quad (34)$$

Note that in above formulation the pair defined by $\theta := (\epsilon, \zeta)$ is viewed as a design parameter vector for the probabilistic certification setting. The same notation $\theta_\epsilon = \epsilon$ and $\theta_\zeta = \zeta$ will also be used to invoke the individual components of this design vector. The above notation enables to formulate the final form of the ϵ -observability formulation:

PROPOSITION 3 (FINAL FORMULATION): The observation target is ϵ -observable on \mathbb{W} with the dead-zone vector ζ if and only if the pair $\theta = (\epsilon, \zeta)$ satisfies the following condition:

$$(\forall w \in \mathbb{W}) \quad g(w, \theta) = 0 \quad (35)$$

The condition (35) is called a **robust constraints satisfaction condition** as the satisfaction of the constraint $g(w, \theta) = 0$ is required for **all** possible realizations of the argument w and this, regardless of its probability of occurrence. This drawback is handled by the probabilistic certification framework leading to the concept of η -almost ϵ -observability which is introduced in the following section.

IV. THE η -ALMOST ϵ -OBSERVABILITY

In order to avoid these difficulties and to come with a realistic assessment of the observability, the following less stringent concept of observability is introduced:

DEFINITION 5: [η -Almost ϵ -observability] Given a predefined observation-target variable z , given a small $\eta \in (0, 1)$ and assuming some probability distribution \mathcal{W} governing the dispersion of the context parameter w we say that the observation target z is η -almost ϵ -observable if and only if there is a design parameter $\theta = (\epsilon, \zeta)$ such that the following inequality holds true:

$$\Pr_{\mathcal{W}} \left[g(\cdot, \theta) \neq 0 \right] \leq \eta \quad (36)$$

meaning that the ϵ -observability condition (35) is satisfied with a high probability $(1-\eta)$.

Although the formulation of (36) is less stringent than (35), it is still difficult to manipulate. That is why probabilistic certification approaches have been developed [3], [9] which

introduces a SECOND STEP consisting in approximating (36) via numerical averaging over a high number $N_s(\eta, \delta, m)$ of sampled realizations of the stochastic variable w . This number $N_s(\eta, \delta, m)$ depends on the previously introduced precision η , a parameter $\delta \in (0, 1)$ representing the confidence with which [namely with probability $1-\delta$], the certification result can be stated and the number of accepted failures among the drawn $N_s(\eta, \delta, m)$ samples, namely:

$$\sum_{i=1}^{N_s(\delta, \eta, m)} g(w^{(i)}, \theta) \leq m \quad (37)$$

Based on the above notation, the probabilistic certification of the ϵ -observability concept is given by the following proposition [2]:

PROPOSITION 4 (CERTIFICATION OF ϵ -OBSERVABILITY):

Consider the following setting's components:

- A given discrete set $\Theta \subset \mathbb{R}_+^2$ containing n_Θ values of the design parameter $\theta := (\epsilon, \zeta)$.
- A certification confidence parameter $\delta \in (0, 1)$
- A certification precision parameter $\eta \in (0, 1)$
- A maximum number of failures $m \in \mathbb{N}_*$
- An integer N_s satisfying:

$$N_s \geq \frac{1}{\eta} \left[m + \ln\left(\frac{n_\Theta}{\delta}\right) + (2m \ln\left(\frac{n_\Theta}{\delta}\right))^{\frac{1}{2}} \right] \quad (38)$$

- N_s realizations $\{w^{(i)}\}_{i=1}^{N_s}$ of the vector w drawn randomly according to \mathcal{W} inside \mathbb{W} .

If there is an element $\theta^* = (\epsilon^*, \zeta^*) \in \Theta$ such that the following inequality holds true:

$$\sum_{i=1}^{N_s} g(w^{(i)}, \theta^*) \leq m \quad (39)$$

Then the condition

$$\Pr_{\mathcal{W}} \left(g(\cdot, \theta^*) \neq 0 \right) \leq \eta \quad (40)$$

is satisfied with a probability greater than $1-\delta$. In other words, if the condition (39) holds for N_s satisfying (38) then the observation target is η -almost ϵ^* -observable in the sense of Definition 5.

Based on the above discussion, the randomized optimization framework amounts at solving:

$$\min_{\theta \in \Theta} \left[\text{cost}(\theta) \right] \quad \text{under} \quad \sum_{i=1}^{N_s} g(w^{(i)}, \theta) \leq m \quad (41)$$

where N_s is defined by (38). In our case, the cost function in (41) that is defined in terms of the decision variable $\theta := (\epsilon, \zeta)$ is obviously given by: $\text{cost}(\theta) := \epsilon$ since the objective is to get a certification results with the lowest state estimation error on the observation target variable while ζ is simply a design parameter of the MHE. Before we examine a specific illustrative example, some general comments and discussion is proposed relative to different aspects of the implementation.

A. The design set

For each considered configurations of sampling statistics, N_s samples are drawn with N_s satisfying the inequality (38). The probabilistic certification requires finding $\theta^* = (\epsilon^*, \zeta^*)$ such that the inequality (39) holds over the set of sampled $w^{(i)}$, except at most m instances, where θ^* belongs to a before hand defined discrete set Θ of cardinality n_Θ . In what follows the structure of the set of design parameter Θ is taken of the form:

$$\Theta := \mathbb{L}(\underline{\sigma}_\epsilon, \bar{\sigma}_\epsilon, n_\epsilon) \times \mathbb{L}(\underline{\sigma}_\zeta, \bar{\sigma}_\zeta, n_\zeta) \quad (42)$$

where $\mathbb{L}(\underline{\sigma}, \bar{\sigma}, n)$ is the set of n logarithmically uniformly spaced numbers, namely¹:

$$\mathbb{L}(\underline{\sigma}, \bar{\sigma}, n) = \text{logspace}(\underline{\sigma}, \bar{\sigma}, n) \\ := \left\{ 10^{r_i} \mid r_i = \underline{\sigma} + \frac{(\bar{\sigma} - \underline{\sigma})i}{n-1} \quad i \in \{0, \dots, n-1\} \right\}$$

This obviously leads to a cardinality $n_\Theta = n_\epsilon n_\zeta$

B. Implementation and complexity analysis

It is shown shortly that the specific observability problem leads to a specific complexity analysis but in a more general settings of (41), since we are using a discrete set Θ of admissible values of θ , the worst case analysis involves a number of simulations that is equal to $N_s n_\Theta = N_s n_\epsilon n_\zeta$ simulations which corresponds to an exhaustive search. For the specific certification problem of almost ϵ -observability, it turns out that the structure of the constraint function $g(w, \theta)$ takes the following form $g(w, \theta) = G(e(w), \theta)$ where e is the output prediction error before dead-zone clipping while G is a very cheap map that mainly consists in clipping followed by conditional summation over the resulting profiles. Consequently, an exhaustive search is more affordable than in the general case since the computation consists in three successive steps which are:

- 1) Simulating the N_s scenarios using the set $\{w^{(i)}\}_{i=1}^{N_s}$ sampled instances to get the set

$$\{e^{(i)} := e(w^{(i)})\}_{i=1}^{N_s}$$

- 2) Computing for each $\theta \in \Theta$ the resulting cost and constraints for each $e^{(i)}$ using $G(e^{(i)}, \theta)$
- 3) Select among all values of θ satisfying the constraints the one that corresponds to the lowest $\epsilon = \theta_1$.

The whole certification framework is summarized in Algorithm 1.

V. ILLUSTRATIVE EXAMPLE

Consider the example of the nonlinear continuous stirred-tank reactor with parallel reaction $R \rightarrow (P_1, P_2)$ [11]: that can be described by the following set of dimensionless energy and material balances:

$$\dot{x}_1 = 1 - p_1 x_1^2 e^{-1/x_3} - p_2 e^{-p_3/x_3} - x_1 \quad (43)$$

$$\dot{x}_2 = p_1 x - 1^2 e^{-1/x_3} - x_2 \quad (44)$$

$$\dot{x}_3 = u - x_3 \quad (45)$$

¹As an example, $\mathbb{L}(-2, 0, 5) \approx \{0.01, 0.032, 0.1, 0.32, 1.\}$

Algorithm 1 The certification algorithm

1: **Given:**

- The maps f and h ▷ [see (1) and (2)]
- η, δ, m ▷ Certification parameters [see (38)]
- $\mathbb{X}, \mathbb{P}, \mathbb{U}, \mathbb{V}$ ▷ Working sets
- $\underline{\sigma}_\epsilon, \bar{\sigma}_\epsilon, \underline{\sigma}_\zeta, \bar{\sigma}_\zeta, n_\epsilon, n_\zeta$ ▷ Design Set parameters (42)
- Sampling rule \mathcal{W} on $\mathbb{W} = \mathbb{Q} \times \mathbb{X} \times \mathbb{P}$

2: $n_\Theta \leftarrow n_\epsilon \cdot n_\zeta$ ▷ Cardinality of the design set Θ

3: Compute $\Theta := \{\theta^{(j)}\}_{j=1}^{n_\Theta}$ by (42) in alphabetic order

4: $N_s \leftarrow N_s(\delta, \eta, m, n_\Theta)$ by (38) ▷ Number of scenarios

5: Generate N_s scenarios inputs $\{w^{(i)}\}_{i=1}^{N_s}$ using \mathcal{W}

6: **for** $i \in \{1, \dots, N_s\}$ **do**

7: Compute $e^{(i)}$ ▷ non clipped error for $w^{(i)}$

8: **end for**

9: success \leftarrow False ▷ No solution θ^* found yet

10: **for** $j \in \{1, \dots, n_\Theta\}$ **do**

11: nb_failures $\leftarrow \sum_{i=1}^{N_s} g(w^{(i)}, \theta^{(j)})$

12: **if** nb_failures $\leq m$ **then**

13: success \leftarrow True ▷ Optimal solution found

14: $\theta^* \leftarrow \theta^{(j)}$. **Break.** ▷ Stop the loop

15: **end if**

16: **end for**

17: **Output:** **If** success **return** θ^*

where x_1 and x_2 represent the concentrations of R and P_1 while x_3 stands for the temperature of the mixture in the reactor. P_2 represents the waste product. This reactor is controlled by the manipulated variable $u \in \mathbb{U} := [0.049, 0.449]$ (which represents the reaction temperature) in order to maximize the production of $x_2 = P_1$. Note that the above dynamics involves $n_p = 3$ parameters p_1, p_2 and p_3 whose nominal values are commonly considered to be $p^{nom} = (10^4, 4 \times 10^2, 0.55)$. It is assumed that x_2 is measured together with u while x_1 and x_3 has to be estimated by the observer. The above settings is applied with the following definition of the subset $\mathbb{X} \subset \mathbb{R}^3$:

$$\mathbb{X} := [0, 0.6] \times [0, 0.3] \times [0.05, 0.2]$$

which contains realistic evolutions of the state during realistic operational context. The discrete design set Θ containing the candidate values of the pair $(\epsilon, \zeta) \in \{10^{-4}, \dots, 1\} \times \{10^{-4}, \dots, 10^{-2}\}$ that are uniformly distributed on a logarithmic scale leading to a cardinality $n_\Theta = 200$.

In order to completely define the statistics \mathcal{W} , we need to define the sampling rules of all the sets involved in the definition of \mathbb{W} , namely, $\mathbb{X}, \mathbb{P}, \mathbb{U}$ and \mathbb{V} . The choices used in the numerical investigation of this paper are defined as follows: Regarding the state, a uniform distribution over the hypercube \mathbb{X} . As for \mathbb{P} , two possibilities are investigated, namely: A uniform distribution over \mathbb{P} or a Gaussian distribution around the nominal value. Two possibilities are also investigated regarding the control sequences set \mathbb{U}^N : A uniform distribution over \mathbb{U}^N or a truncated Fourier series projected into the admissible set. Finally, the noise profile realizations have been generated using the uniform

distribution over $[-\bar{\nu}, +\bar{\nu}]$. Three observation-targets are investigated which are $z_1 = x, z_2 = x_1$ or $z_3 = x_3$.

A. Results

Different aspects are successively examined in terms of their impact on the certification results. The results are shown through data frames in which the signification of the columns are as follows: **eps1, eps2, eps3**: The certified reconstruction precision ϵ on z_1, z_2 and z_3 respectively. **zeta1, zeta2, zeta3**: The optimal computed dead-zone sizes. More precisely (**eps1, zeta1**), (**eps2, zeta2**) and (**eps3, zeta3**): The solutions of (41) when the observation-target variable is z_1, z_2 or z_3 respectively. **N**: The observation horizon. **noise**: The noise level $\bar{\nu}$ mentioned above. **rho**: The size ρ of the hyper-box as invoked in when the uniform distribution of the parameters is used (This corresponds to the value *uniform* in the column entitled **p_mode**).**std_p**: used in the gaussian distribution of the parameters around the nominal values. (This corresponds to the value *gaussian* in the column entitled **p_mode**). **u_mode**. The type of input used in the certification (can be *Fourier* or *rand*).

Impact of the observation horizon. Figure 1 shows that one needs to use $N = 20$ in order to achieve the certification with the lowest values $\epsilon = 10^{-4}$ considered in the design set Θ over the three observation-target variables $z_i, i = 1, 2, 3$. Otherwise, indistinguishability might occur.

Impact of the measurement noise. Starting from the last setting of Figure 1, the noise level is increased from 0.001 to 0.003. This leads to a sensitive degradation in the certifiable reconstruction precision. Figure 2 shows that by increasing the observation horizon up to $N = 100$ it is possible to recover the levels of precision of the second setting of Figure 1 which was achievable with $N = 10$ and the previous level (0.001) of the noise. This clearly shows that the higher the noise is the longer the observation horizon should be to achieve the same level of certifiable reconstruction precision.

Impact of the parametric uncertainty level and statistics.

The comparison between the first two lines of Figure 3 shows that the configuration with uniform distribution of the parameter vector with a level $\rho = 0.05$ is more inconvenient to certifiable reconstruction precision than the gaussian distribution with $std = 0.2$.

Impact of the input. In the previous results, the control profile was systematically taken to be a randomly sampled truncated Fourier series. Figure 4 shows the results for uniformly randomly generated profiles inside the admissible set $[0.049, 0.449]$. It can be observed by comparing the first lines of Figures 3 and 4 that random profiles seem to enhance the observability at least for the setting that is common to these two lines.

Impact of the certification precision parameter. Figure 5 shows how the results change when the targeted precision

eps1	eps2	eps3	zeta1	zeta2	zeta3	N	noise	rho	std_p	u_mod	p_mod	eta	delta	m
0.379269	0.379269	0.054556	0.000774	0.000774	0.000774	5	0.001000	0.050000	-	Fourier	uniform	0.010000	0.001000	10
0.233572	0.143845	0.054556	0.000774	0.000774	0.000774	10	0.001000	0.050000	-	Fourier	uniform	0.010000	0.001000	10
0.000100	0.000100	0.000100	0.000464	0.000464	0.000464	20	0.001000	0.050000	-	Fourier	uniform	0.010000	0.001000	10

Fig. 1. Impact of the observation horizon on observability.

eps1	eps2	eps3	zeta1	zeta2	zeta3	N	noise	rho	std_p	u_mod	p_mod	eta	delta	m
0.000100	0.000100	0.000100	0.000464	0.000464	0.000464	20	0.001000	0.050000	-	Fourier	uniform	0.010000	0.001000	10
0.379269	0.379269	0.088587	0.001292	0.001292	0.001292	20	0.003000	0.050000	-	Fourier	uniform	0.010000	0.001000	10
0.233572	0.143845	0.054556	0.000774	0.000774	0.000774	100	0.003000	0.050000	-	Fourier	uniform	0.010000	0.001000	10

Fig. 2. Impact of the measurement noise level on observability.

eps1	eps2	eps3	zeta1	zeta2	zeta3	N	noise	rho	std_p	u_mod	p_mod	eta	delta	m
0.233572	0.088587	0.033598	0.000464	0.000464	0.000464	20	0.001000	0.050000	-	Fourier	uniform	0.010000	0.001000	10
0.000100	0.000100	0.000100	0.000464	0.000464	0.000464	20	0.001000	-	0.200000	Fourier	gaussian	0.010000	0.001000	10
0.088587	0.033598	0.020691	0.000464	0.000464	0.000464	20	0.001000	-	0.300000	Fourier	gaussian	0.010000	0.001000	10

Fig. 3. Impact of the parameter uncertainties level and distribution on observability.

eps1	eps2	eps3	zeta1	zeta2	zeta3	N	noise	rho	std_p	u_mod	p_mod	eta	delta	m
0.000100	0.000100	0.000100	0.000464	0.000464	0.000464	20	0.001000	0.050000	-	rand	uniform	0.010000	0.001000	10
0.143845	0.033598	0.012743	0.000464	0.000464	0.000464	20	0.001000	0.050000	0.200000	rand	gaussian	0.010000	0.001000	10
0.233572	0.088587	0.054556	0.000464	0.000464	0.000464	20	0.001000	0.050000	0.300000	rand	gaussian	0.010000	0.001000	10
0.000100	0.000100	0.000100	0.000278	0.000278	0.000278	50	0.001000	0.050000	0.300000	rand	gaussian	0.010000	0.001000	10
0.000100	0.000100	0.000100	0.000278	0.000278	0.000278	50	0.001000	0.050000	0.300000	Fourier	gaussian	0.010000	0.001000	10

Fig. 4. Checking observability with random input.

η of the probabilistic certification is degraded. As expected, one can see that under the settings of this figure (the same for all values of η), a certification precision of $\eta = 10^{-3}$ seems to lead to a certifiable upper bounds on the estimation errors on the observation-target variables which are quite high (roughly useless given the definition of the set \mathbb{X}). This is obviously due to the high level of parameter dispersion $std = 0.3$. The certifiable reconstruction error decreases when η is increased meaning that the part of the pairs over which the reported reconstruction error are guaranteed is a smaller set of admissible pairs. For instance, the last line of Figure 5 indicates that up to 5% of the samples correspond to the presence of indistinguishable pairs. For the remaining 95% of the cases, an almost zero reconstruction error can be certified (provided that the optimization problem is correctly solved).

B. Computation times and scalability

As it has been mentioned in Section IV-B, the computation time is the sum of the CPU times needed for two successive tasks, namely:

- **dt1**: the time needed to generate the scenarios
- **dt2**: the time needed to find the optimal design parameter.

These times are mainly impacted by the following parameters:

- 1) The number of scenarios to be simulated which linearly affects **dt1** but not **dt2**.
- 2) The time needed to simulate a single scenario which is for

a given system depends on the observation horizon N that linearly impacts the CPU time **dt1** when fixed step integration scheme is used to integrate ODE's models.

3) The search algorithm that is adopted to find the optimal design parameter. In the above computation, a simple alphabetic search is adopted with increasing $\epsilon = \theta_1$.

Note that the formulae (38) does not depend on the state or the parameter vector's dimension. This means that the size of the system and the number of its uncertain parameters do not lead to an exponential increase in the computation time, only the simulation time would affect linearly the computation time of the certification scheme. Note also that both n_Θ and δ appears logarithmically in the expression (38). This same expression shows clearly that increasing the confidence of the certification by reducing δ linearly increases the number of scenarios.

As far as the example is concerned, Figure 6 shows the computation times for the certification scheme as a function of the observation horizon or as a function of the certification precision η . The certification confidence parameter $\delta = 0.001$ and the number of scenarios with failures $m = 10$ are used. Note how the CPU time for the second task (optimizing the design parameter) decreases with the observation horizon as the number of values to be visited (in alphabetic order) is reduced because of the observability

eps1	eps2	eps3	zeta1	zeta2	zeta3	N	noise	rho	std_p	u_mod	p_mod	eta	delta	m
0.615848	0.615848	0.143845	0.000464	0.000464	0.000464	20	0.001000	0.050000	0.300000	rand	gaussian	0.001000	0.001000	10
0.379269	0.379269	0.088587	0.000464	0.000464	0.000464	20	0.001000	0.050000	0.300000	rand	gaussian	0.005000	0.001000	10
0.233572	0.143845	0.054556	0.000464	0.000464	0.000464	20	0.001000	0.050000	0.300000	rand	gaussian	0.010000	0.001000	10
0.000100	0.000100	0.000100	0.000464	0.000464	0.000464	20	0.001000	0.050000	0.300000	rand	gaussian	0.050000	0.001000	10

Fig. 5. Impact of the precision η required in the probabilistic certification.

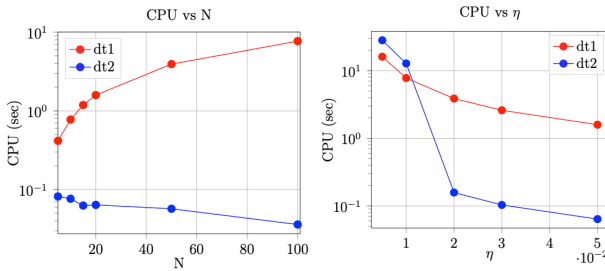


Fig. 6. Computation times using $\delta = 0.001$, $m = 10$. (left) CPU (sec) versus the observation horizon N for $\eta = 0.01$. (right) CPU (sec) versus η for an observation horizon of $N = 20$.

gained by the use of higher observation horizon.

VI. CONCLUSION AND FUTURE POSSIBLE INVESTIGATION

This paper proposes a general scalable scheme for the analysis of observability and parameter reconstruction in the context of nonlinear dynamical systems that are subjected to parametric uncertainties and measurement noise. The observability is taken in a more general sense than the standard (extended) observability commonly used in the sense that it is the possibility and the precision to which it is possible to reconstruct specific expressions of the state/parameter that can be investigated by the proposed scheme. A possible continuation of the present work concerns the investigation of the regions of the space of state and parameters where the certification constraint does not hold leading to degraded certification results. This can be an important step in the analysis since these regions of the space might have been wrongly included while they are obviously to be excluded by the very definition of the operational space of the system. In such cases, these regions should be removed and the computation re-done in order to come out with more consistent results. Another undergoing work consists in applying the proposed scheme to standard models that are widely used in control and analysis of biological systems (diabetes, cancer, HIV, pandemics propagation models, etc) since these systems are by nature defined up to the knowledge of a high number of highly uncertain parameters.

REFERENCES

- [1] M. Alami, J. S. Welsh, and G. C. Goodwin. Redundancy versus multiple starting points in nonlinear system related inverse problems. *Automatica*, 45(4):1052 – 1057, 2009.
- [2] T. Alamo, J.M. Bravo, and E.F. Camacho. Guaranteed state estimation by zonotopes. *Automatica*, 41(6):1035 – 1043, 2005.
- [3] T. Alamo, R. Tempo, and E.F. Camacho. Randomized strategies for probabilistic solutions of uncertain feasibility and optimization problems. *Automatic Control, IEEE Transactions on*, 54(11):2545–2559, Nov 2009.
- [4] A. Alessandri, M. Baglietto, G. Battistelli, and V. Zavala. Advances in moving-horizon estimation for nonlinear systems. In *Proceedings of the IEEE Conference on Decision and Control*, pages 5681–5688, 2010.
- [5] A. Alessandri, M. Baglietto, and G. Battistelli. Moving-horizon state estimation for nonlinear discrete-time systems: New stability results and approximation schemes. *Automatica*, 44:1753–1765, 2008.
- [6] Gillis J. Horn G. et al Andersson, J.A.E. Casadi: a software framework for nonlinear optimization and optimal control. *Journal of Theoretical Biology*, 384:59 – 69, 2015.
- [7] D. Astolfi and L. Marconi. A high-gain nonlinear observer with limited gain power. *IEEE Transactions on Automatic Control*, 60(11):3059–3064, 2015.
- [8] G. Besançon. *Nonlinear Observers and Applications*. Lecture Notes on Control and Information Science (LNCIS). Springer-Verlag, 2007.
- [9] G. C. Calafiore and M. C. Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2006.
- [10] S. Drakunov and V. Utkin. Sliding mode observers. tutorial. In *Proceedings of 1995 34th IEEE Conference on Decision and Control*, volume 4, pages 3376–3378 vol.4, 1995.
- [11] Bailey J. E., F. J. M. Horn, and R. C. Lin. Cyclic operation of reaction systems: effect of heat and mass transfer resistance. *AIChE Journal*, 17(4):818–825, 1971.
- [12] Denis Efimov, Tarek Raussi, Stanislav Chebotarev, and Ali Zolghadri. Interval state observer for nonlinear time varying systems. *Automatica*, 49(1):200 – 205, 2013.
- [13] R. Hermann and A. Krener. Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, 22(5):728–740, 1977.
- [14] L. Ji, J. B. Rawlings, W. Hu, A. Wynn, and M. Diehl. Robust stability of movinf-horizon estimation under bounded disturbances. *IEEE Transactions on Automatic Control*, 61(11):3509–3514, 2016.
- [15] Costas Kravaris. Functional observers for nonlinear systems. *IFAC-PapersOnLine*, 49(18):505 – 510, 2016. 10th IFAC Symposium on Nonlinear Control Systems NOLCOS 2016.
- [16] Ping Li and James Lam. Positive state-bounding observer for positive interval continuous-time systems with time delay. *International Journal of Robust and Nonlinear Control*, 22(11):1244–1257, 2012.
- [17] Nacim Meslem, Nacim Ramdani, and Yves Candau. Interval observers for uncertain nonlinear systems. application to bioreactors. *IFAC Proceedings Volumes*, 41(2):9667 – 9672, 2008.
- [18] H. Michalska and D. Q. Mayne. Moving horizon observers and observer-based control. *IEEE Transactions on Automatic Control*, 40(6):995–1006, 1995.
- [19] M. A. Muller. Nonlinear moving horizon estimation in the presence of bounded disturbances. *Automatica*, 79:306–314, 2017.
- [20] Kenneth R. Muske and James B. Rawlings. *Nonlinear Moving Horizon State Estimation*, pages 349–365. Springer Netherlands, Dordrecht, 1995.
- [21] Christopher V. Rao and James B. Rawlings. Nonlinear moving horizon state estimation. In Frank Allgöwer and Alex Zheng, editors, *Nonlinear Model Predictive Control*, pages 45–69. Basel, 2000. Birkhäuser Basel.
- [22] J. B. Rawlings and L. Li. Optimization-based state estimation: Current status and some new results. *Journal of Process Control*, 22:1439–1444, 2012.
- [23] P.S. Teh and H. Trinh. Design of unknown input functional observers for nonlinear systems with application to fault diagnosis. *Journal of Process Control*, 23(8):1169 – 1184, 2013.
- [24] Hieu Trinh and Tyrone Fernando. *Linear Functional Observers*, pages 49–72. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.