Robust stability of moving horizon estimation for nonlinear systems with bounded disturbances using adaptive arrival cost

Nestor Deniz[1], *Marina Murillo*[1], *Guido Sanchez*[1], *Leonardo Giovanini*[1]

[1] *Research institute for signals, systems and computational intelligence, Ciudad Universitaria UNL, Ruta Nac. No 168, km 472.4, FICH, 4to Piso (3000) Santa Fe - Argentina*

* *E-mail: ndeniz@sinc.unl.edu.ar*

Abstract: The robust stability and convergence to the true state of a moving horizon estimator based on an adaptive arrival cost are established for nonlinear detectable systems in this paper. Robust global asymptotic stability is shown for the case of non-vanishing bounded disturbances, whereas the convergence to the true state is proved for the case of vanishing disturbances. Two simulations were made to show the estimator behaviour under different operational conditions and to compare it with the state of the art of estimation methods.

1 Introduction

State estimation plays a fundamental role in feedback control, system monitoring, and system optimization because noisy measurements is the only information available from the system. Several methods have been developed for accomplishing such task (see [1, 2], among others). All these methods have been developed upon assumptions on the knowledge of noises and the system model as well as the absence of constraints.

In practice, these assumptions are not easily satisfied, and research efforts were focused on approaches that do not rely on such requirements (see [3–5], among others). For example, H_∞ filters are designed minimizing the H_∞ norm of the mapping between disturbances and estimation error. Estimators based on least-square estimation problems have been presented in [3, 6] and [7]. Both approaches are based on the adequate selection of the uncertainty model, resting on the available information of the system, instead of relying on statistical assumptions on noises. In a similar way, robust estimation algorithms based on $\min$-$\max$ robust filtering, set-valued estimation and guaranteed cost paradigm, have attracted the attention of the research community (see [4, 8]).

Building on the success of moving horizon control, moving horizon estimation (*MHE*) has attracted the attentionof researchers since the pioneering work of Jazwinski [9] (see also [10–12]). The interest in such estimation methods stems from the possibility of dealing with a limited amount of data and the ability to incorporate constraints. In recent years, both theoretical properties of several *MHE* schemes, including efficient computational methods, have been studied (see [13–18]).

In recent years several results on robust stability and convergence properties have been obtained, advancing from idealistic assumptions (observability and no disturbances) to realistic situations (detectability and bounded disturbances). For nonlinear observable systems, [12] established the asymptotic stability of the estimation error for the least-square cost function. Furthermore, for convergent disturbances, the estimation error is also convergent [19–21]. Alessandri et al. proposed an estimation scheme, based on a least-square cost function of estimation residuals, that guaranteed the boundedness of estimation error for observable systems subject to bounded additive disturbances [14, 22]. The review made by Rawlings and Li [20] provides a general view of the problem relying on incremental input/output-to-state stability (*i-IOSS*, see Definition 2.1 in Section 2) for detectability [23] and robust global asymptotic stability (*RGAS*) for robust stability of the state estimator. This work reveal two major challenges in the field: *i*) conditions and a proof of *RGAS* for a *MHE* subject to bounded disturbances and *ii*) the development of computational efficient *MHE* algorithms. Hu et al. [24] identifies a broad class of cost functions that ensures *RGAS* of full information estimators (*FIE*s). The implication of *RGAS* on *MHE* was further investigated in [25] based on the results of [26]. Moreover, [25] showed *RGAS* and convergence of estimation error in case of bounded or vanishing disturbances, respectively. In these works, the least-square cost function was modified by adding a $\max$-term to guarantee stability. For a particular choice of weights of the cost function, these results were extended to least-squares type [25]. Finally, the necessary assumptions on the cost function were generalized in [27].

The results described for nonlinear detectable systems subject to bounded disturbances show a number of drawbacks. In particular, the disturbances gains obtained in [25] are conservative, and they depend on the estimation horizon. The same happens with the estimate on the minimal estimation horizon in the case of *MHE*. These estimates depend on a priori bounds of worst case disturbances. One attempt to solve these flaws is presented in [28], where the authors introduce a novel formulation of the cost function that allows ensuring *robust global exponential stability* of the estimation error under a suitable exponential detectability. Using this idea, they obtain improved estimates for the disturbance gains and the minimal estimation horizon.

This paper introduces the *RGAS* and convergence analysis for a *MHE* estimator based on adaptive arrival cost proposed in [18] for the practical case of nonlinear detectable systems subject to bounded disturbances. This formulation allows us to overcome several drawbacks in the existing literature, obtaining improved estimates for the disturbance gains and the minimal estimation horizon and providing stability proof for both *FIE* and *MHE*. To establish robust stability properties for *MHE*, it is essential the adequate selection of the arrival cost employed by the cost function. In various *MHE* schemes, the necessary assumptions of the arrival cost are difficult to verify ([12, 19]), while in others, they can be verified a priori [25]. In the *MHE* scheme analyzed in this work, the assumption on the arrival cost can be satisfied a priori by design. Furthermore, the disturbances gains become uniform (they are valid independent of estimation horizon N), allowing to extend the stability analysis to *FIE* estimators with least-square cost functions.

The rest of the paper is organized as follows: Section 2 introduces the notation, definitions and properties that will be used through the paper. Section 3 presents the main result and shows its connections with previous stability analysis. Section 4 discusses two examples previously used in the literature with the purpose of illustrating the concepts and showing the difference with others *MHE* algorithms. Finally, Section 5 presents conclusions.

2 Preliminaries and setup

2.1 Notation

Let $\mathbb{Z}_{[a,b]}$ denotes the set of integers in the interval $[a,b]$ and $\mathbb{Z}_{\geq a}$ denotes the set of integers greater or equal to a. Boldface symbols denote finite sequences $\boldsymbol{w}_{[k_1,k_2]} := \{w_{k_1},\ldots,w_{k_2}\}$ for some $k_1,k_2 \in \mathbb{Z}_{\geq 0}$ and $k_1 < k_2$, and infinite sequences $\boldsymbol{w} := \{w_k, w_{k+1},\ldots\}$, $k \in \mathbb{Z}_{\geq 0}$, respectively. We denote $\hat{x}_{j|k}$ as the estimation of x_j at time $k \in \mathbb{Z}_{\geq 0}$. By $|x|$ we denote the Euclidean norm of a vector x. Let $\|\boldsymbol{x}\| := \sup_{k\in\mathbb{Z}_{\geq 0}} |x_k|$ denote the supreme norm of the sequence $\boldsymbol{x}$ and $\|\boldsymbol{x}\|_{[a,b]} := \sup_{k\in\mathbb{Z}_{[a,b]}} |x_k|$. For short, we will use simply $\|\boldsymbol{x}\|$ without indices where it is clear from the context. A function $\gamma : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class $\mathcal{K}$ if γ is continuous, strictly increasing and $\gamma(0) = 0$. If γ is also unbounded, it is of class $\mathcal{K}_\infty$. A function $\zeta : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class $\mathcal{L}$ if $\zeta(k)$ is non increasing and $\lim_{k\to\infty} \zeta(k) = 0$. A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{Z}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class $\mathcal{KL}$ if $\beta(\cdot,k)$ is of class $\mathcal{K}$ for each fixed $k \in \mathbb{Z}_{\geq 0}$, and $\beta(r,\cdot)$ of class $\mathcal{L}$ for each fixed $r \in \mathbb{R}_{\geq 0}$.

The following inequalities hold for all $\beta \in \mathcal{KL}$, $\gamma \in \mathcal{K}$ and $a_j \in \mathbb{R}_{\geq 0}$ with $j \in \mathbb{Z}_{[1,n]}$

$$\gamma\left(\sum_{j=1}^{n} a_i\right) \leq \sum_{j=1}^{n} \gamma(na_i), \quad \beta\left(\sum_{j=1}^{n} a_i, k\right) \leq \sum_{j=1}^{n} \beta(na_i, k). \tag{1}$$

The preceding inequalities hold since $\max\{a_j\}$ is included in the sequence $\{a_1, a_2, \ldots, a_n\}$ and $\mathcal{K}$ functions are non-negative strictly increasing functions.

A sequence $\boldsymbol{w}$ is bounded if $\|\boldsymbol{w}\|$ is finite. The set of bounded sequences $\boldsymbol{w}$ is denoted as $\mathcal{W}(w_{\max}) := \{w : \|\boldsymbol{w}\| \leq w_{\max}\}$ for some $w_{\max} \in \mathbb{R}_{\geq 0}$. Moreover, a bounded infinite sequence $\boldsymbol{w}$ is convergent if $|w_k| \to 0$ as $k \to \infty$. Let us denote the set of convergent sequences $\mathcal{C}$:

$$\mathcal{C}_w := \{\boldsymbol{w} \in \mathcal{W}(w_{\max}) \mid \boldsymbol{w} \text{ is convergent}\}. \tag{2}$$

Analogously, $\mathcal{C}_v$ is defined for the sequence $\boldsymbol{v}$.

2.2 Problem statement

Let us consider the state estimation problem for nonlinear discrete time systems of the form

$$\begin{aligned} x_{k+1} &= f(x_k, w_k), \\ y_k &= h(x_k) + v_k, \end{aligned} \tag{3}$$

where $x_k \in \mathcal{X} \subseteq \mathbb{R}^{n_x}$, $w_k \in \mathcal{W} \subset \mathbb{R}^{n_w}$, $y_k \in \mathcal{Y} \subseteq \mathbb{R}^{n_y}$ and $v_k \in \mathcal{V} \subset \mathbb{R}^{n_v}$ are the state, process disturbance, output measurement and measurement disturbance vectors, respectively. The process w_k and measurement v_k disturbances are unknown but bounded, i.e, $\boldsymbol{w} \in \mathcal{W}(w_{\max})$ and $\boldsymbol{v} \in \mathcal{V}(v_{\max})$ for some $w_{\max}, v_{\max} \in \mathbb{R}_{\geq 0}$. $\mathcal{X}, \mathcal{Y}, \mathcal{W}$ and $\mathcal{V}$ are compact and convex sets with the null vector $\mathbf{0}$ belonging to them. In the following we assume that $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_w} \to \mathbb{R}^{n_x}$ is locally Lipschitz on its arguments and $h : \mathbb{R}^{n_x} \to \mathbb{R}^{n_y}$ is continuous. The solution of system (3) at time k is denoted by $x(k; x_0, \boldsymbol{w})$, with initial condition x_0 and process disturbance sequence $\boldsymbol{w}$. The initial condition x_0 is unknown, but a prior knowledge $\bar{x}_0$ is available and its error is bounded

$$\bar{x}_0 \in \mathcal{X}_0 := \{\bar{x}_0 : |x_0 - \bar{x}_0| \leq e_{\max}\}, \quad \mathcal{X}_0 \subseteq \mathcal{X}. \tag{4}$$

The solution of an estimation problem aims to find an estimate $\hat{x}_{k|k}$ of the current state x_k at time k by minimizing a performance metric Ψ. At each sampling time k, given the previous N measurements $\boldsymbol{y}_{[k-N,k]}$, the following optimization problem is solved

$$\begin{gathered} \min_{\hat{x}_{k-N|k}, \hat{\boldsymbol{w}}} \Psi := \Gamma_{k-N|k}(\chi) + \sum_{j=k-N}^{k} \ell\left(\hat{w}_{j|k}, \hat{v}_{j|k}\right) \\ \text{s.t.} \begin{cases} \chi = \hat{x}_{k-N|k} - \bar{x}_{k-N}, \\ \hat{x}_{j+1|k} = f\left(\hat{x}_{j|k}, \hat{w}_{j|k}\right), \\ y_j = h\left(\hat{x}_{j|k}\right) + \hat{v}_{j|k}, \\ \hat{x}_{j|k} \in \mathcal{X},\ \hat{w}_{j|k} \in \mathcal{W},\ \hat{v}_{j|k} \in \mathcal{V}, \end{cases} \end{gathered} \tag{5}$$

where $\hat{x}_{j|k}$ is the estimate of the state at time j and $\hat{w}_{j|k}$ is the estimate of the process noise at time j, estimated at time k based on measurements $\boldsymbol{y}_{k-N,k}$ available at time k. The prior estimate $\hat{x}_{k-N|k}$ and the process noise $\hat{\boldsymbol{w}}_{[k-N|k]}$ are the optimization variables. The stage cost $\ell\left(\hat{w}_{j|k}, \hat{v}_{j|k}\right)$ penalizes the estimated process disturbance $\hat{w}_{j|k}$ and the estimation residuals $\hat{v}_{j|k} = y_j - h\left(\hat{x}_{j|k}\right)$, while the arrival cost $\Gamma_{k-N|k}(\chi)$ penalizes the prior estimate $\hat{x}_{k-N|k}$. $\ell(\cdot,\cdot)$ and $\Gamma_{k-N|k}(\cdot)$ and their parameters, allows to ensure the robust stability of the estimator [25]. When the estimation window is not full, $k < N$, problem (5) is reformulated and solved as a full information estimator (*FIE*)

$$\begin{gathered} \min_{\hat{x}_{0|k}, \hat{\boldsymbol{w}}} \Psi := \Gamma_{0|k}(\chi_0) + \sum_{j=1}^{k} \ell\left(\hat{w}_{j|k}, \hat{v}_{j|k}\right) \\ \text{s.t.} \begin{cases} \chi_0 = \hat{x}_{0|k} - \bar{x}_0, \\ \hat{x}_{j+1|k} = f\left(\hat{x}_{j|k}, \hat{w}_{j|k}\right), \\ y_j = h\left(\hat{x}_{j|k}\right) + \hat{v}_{j|k}, \\ \hat{x}_{j|k} \in \mathcal{X},\ \hat{w}_{j|k} \in \mathcal{W},\ \hat{v}_{j|k} \in \mathcal{V}, \end{cases} \end{gathered} \tag{6}$$

as k increases this problem becomes (5) $\forall k \in \mathbb{Z}_{\geq N}$.

In previous works, the robust stability of *MHE* has been achieved by modifying the standard least-square cost function through the inclusion of a max-term [25, 26] or by a suitable choice of the cost's function parameters [25]. Another way to solve this problem is combining a suitable choice of $\ell\left(\hat{w}_{j|k}, \hat{v}_{j|k}\right)$ with a time–varying arrival cost of the form

$$\Gamma_{k-N|k}(\chi) = |\,\hat{x}_{k-N|k} - \bar{x}_{k-N}\,|_{P^{-1}_{k-N|k}}, \tag{7}$$

whose parameters $P^{-1}_{k-N|k}$ and $\bar{x}_{k-N}$ are recursively updated using the information available at time k [18]. This way of defining $\Gamma_{k-N|k}$ avoids the introduction of artificial cycling in the estimation process (see [19]). The arrival cost matrix $P_{k-N|k}$ is given by

$$P_{k-N|k} = \begin{cases} \frac{1}{\alpha_k} W_k & \text{if } \frac{1}{\alpha_k}\operatorname{Tr}(W_k) \leq c, \\ W_k & \text{otherwise}, \end{cases} \tag{8}$$

with

$$\alpha_k = 1 - \frac{1}{N_k},$$

$$N_k = \left[1 + \hat{x}^T_{k-N|k-1} P_{k-N-1|k-1} \hat{x}_{k-N|k-1}\right] \frac{\sigma}{|\hat{v}_{k-N|k}|^2},$$

$$\hat{v}_{k-N|k} = y_{k-N} - \hat{y}_{k-N|k},$$

$$W_k = P_{k-N-1|k-1} - \frac{P_{k-N-1|k-1}\hat{x}_{k-N|k-1}\hat{x}^T_{k-N|k-1}}{1+\hat{x}^T_{k-N|k-1}P_{k-N-1|k-1}\hat{x}_{k-N|k-1}P_{k-N-1|k-1}},$$

where σ, c, $\lambda \in R_{>0}$ are tuning parameters, $c > \lambda$, $P_0 = \lambda I_{n\times n}$ and $\sigma \gg \sigma_w$, where σ_w denotes the process noise variance. The prior knowledge of the window $\bar{x}_{k-N}$ is updated using a smoothed estimate [29]

$$\bar{x}_{k-N} = \hat{x}_{k-N|k-1}. \tag{9}$$

The optimization problem (5) can be reformulated in terms of the initial condition $\hat{x}_{0|k}$, the estimated process noises $\hat{\boldsymbol{w}}_{j|k}$ and the measurement noise $\hat{\boldsymbol{v}}_{j|k}$ as follows

$$\min_{\hat{x}_{0|k},\hat{\boldsymbol{w}}} \Psi := \alpha_k^k\, \Gamma_{0|k}\left(\chi_0\right) + \sum_{j=1}^{k-N-1} \alpha_k^{k-j} \ell\left(\hat{w}_{j|k}, \hat{v}_{j|k}\right) + \sum_{j=k-N}^{k-1} \ell\left(\hat{w}_{j|k}, \hat{v}_{j|k}\right) \tag{10}$$

$$\text{s.t.} \begin{cases} \chi_0 = \hat{x}_{0|k} - \bar{x}_0, \\ \hat{x}_{j+1|k} = f\left(\hat{x}_{j|k}, \hat{w}_{j|k}\right), \\ y_j = h\left(\hat{x}_{j|k}\right) + \hat{v}_{j|k}, \\ \hat{x}_{j|k} \in \mathcal{X},\ \hat{w}_{j|k} \in \mathcal{W},\ \hat{v}_{j|k} \in \mathcal{V}. \end{cases}$$

with $\alpha_k \in (0, 1]$. This problem allows to explicitly see the effect of past data on $\hat{x}_{k|k}$. It is easy to see the exponential weighting on $\hat{x}_{0|k}$, $\hat{w}_{j|k}$ and $\hat{v}_{j|k}$ $j \in \mathbb{Z}_{0,k-N-1}$ by $\alpha_k^j \in (0, 1]$ that deemphasizes their effect on $\hat{x}_{k|k}$. In this way, only $\hat{w}_{j|k}$ and $\hat{v}_{j|k}$ $j \in \mathbb{Z}_{k-N,k}$ has full effect on $\hat{x}_{k|k}$.

As α_k change in time as a function of the estimation residual $\hat{v}_{k-N|k}$, modulates weighting based on the information available in $\boldsymbol{y}_{[k-N,k]}$. This way of computing $P_{k-N|k}$ introduces a feedback action between the arrival cost matrix $P_{k-N|k}$ and the estimation residual $\hat{v}_{k-N|k}$. The updating mechanism (8) provides adaptation capabilities that allows the estimator to incorporate the relevant information available in $\boldsymbol{y}_{[0,k]}$ and follow changes in the system (see Figures 3 and 4 of [18]).

The updating mechanism (8) is a time-varying filter whose inputs are $\hat{x}_{k-N|k-1}\hat{x}^T_{k-N|k-1}$ and the initial condition P_0. It generates recursively a real-time estimation of $P_{k-N|k}$ by updating $P_{k-N-1|k-1}$ with an exponential time-averaging of $\hat{x}_{k-N|k-1}\hat{x}^T_{k-N|k-1}$.

Remark 2.1 The sequence $\boldsymbol{P} := \left\{P_{0|0}, \ldots, P_{0|N-1}, P_{k-N|k}\right\}$ $\forall k \in \mathbb{Z}_{\geq 0}$ is positive definite, it is decreasing in norm and it is bounded. The proof of these properties follows similar steps as in [18].

Now, we introduce a definition of detectability for nonlinear systems using a stability definition [30], [31].

Definition 2.1 The system (3) is *incrementally input/output-to-state stable* (i-IOSS) if there exist functions $\beta \in \mathcal{KL}$ and γ_1, $\gamma_2 \in \mathcal{K}$ such that for every initial states z_1, $z_2 \in \mathbb{R}^{n_x}$, any two feasible disturbances sequences $\boldsymbol{w}_1$, $\boldsymbol{w}_2 \in \mathbb{R}^{n_w}$ and output sequences $\boldsymbol{y}_1$, $\boldsymbol{y}_2$, the following inequality holds

$$|x(k, z_1, \boldsymbol{w}_1) - x(k, z_2, \boldsymbol{w}_2)| \leq \max\Big\{\beta\left(|z_1 - z_2|, k\right), \gamma_1\left(\|\boldsymbol{w}_1 - \boldsymbol{w}_2\|_{[0,k-1]}\right), \gamma_2\left(\|h(\boldsymbol{x}_1) - h(\boldsymbol{x}_2)\|_{[0,k-1]}\right)\Big\}, \tag{11}$$

where $x_1 = x(k, z_1, \boldsymbol{w}_1)$ and $x_2 = x(k, z_2, \boldsymbol{w}_2)$. Basically, this definition compares any two system trajectories ($x\left(k; z_1, \boldsymbol{w}_1\right)$ and $x\left(k; z_2, \boldsymbol{w}_2\right)$), verifying the stability of their difference $\forall k \in \mathbb{Z}_{\geq 0}$

$$|x(k, z_1, \boldsymbol{w}_1) - x(k, z_2, \boldsymbol{w}_2)| \leq \eta, \quad \eta \in \mathbb{R}_{\geq 0}. \tag{12}$$

The difference $\eta \in \mathbb{R}_{\geq 0}$ would be stable if and only if the system is detectable and stabilizable. The value of η depends on the difference between the disturbance sequences $\boldsymbol{w}_1, \boldsymbol{w}_2$ and output sequences $h(\boldsymbol{x}_1)$, $h(\boldsymbol{x}_2)$. Since $y = h(x) + v$, then $y_1 = h(x\left(k, z_1, \boldsymbol{w}_1\right)) + \boldsymbol{v}_1$ and $y_2 = h(x\left(k, z_2, \boldsymbol{w}_2\right)) + \boldsymbol{v}_2$. If we take $x\left(k, z_1, \boldsymbol{w}_1\right)$ as the true trajectory of the system and $x\left(k, z_2, \boldsymbol{w}_2\right)$ as the estimated trajectory, then $y_j = h(x_j) + v_j$ and $y_j = h(\hat{x}_{j|k}) + \hat{v}_{j|k}$ and the difference $\|h(\boldsymbol{x}_1) - h(\boldsymbol{x}_2)\|$ becomes $\|\boldsymbol{v}_1 - \boldsymbol{v}_2\|$.

In the following section the updating mechanism (8) and the assumption of ***i-IOSS*** [32] will be used to prove robust stability of the proposed *MHE* in the presence of bounded disturbances and convergence to the true state ($\hat{x}_{k|k} \to x_k$) in the case of convergent disturbances ($w_1, w_2 \in \mathcal{C}_w$, $v_1, v_2 \in \mathcal{C}_v$). Before to proceed to the development of the main results, we state the main properties of system (3) and assumptions about $\Gamma_{k-N}\left(\cdot\right)$ and $\ell\left(\cdot,\cdot\right)$.

Assumption 2.1 The prior weighting $\Gamma_{k-N|k}$ is a continuous function $\Gamma_{k-N|k} : \mathbb{R}^{n_x} \to \mathbb{R}$ lower bounded by $\underline{\gamma}_p : \mathbb{R} \to \mathbb{R}$, $\underline{\gamma}_p \in \mathcal{K}_\infty$ and upper bounded by $\bar{\gamma}_p : \mathbb{R} \to \mathbb{R}$, $\bar{\gamma}_p \in \mathcal{K}_\infty$ such that

$$\underline{\gamma}_p\left(|\chi|\right) \leq \Gamma_{k-N|k}\left(\chi\right) \leq \bar{\gamma}_p\left(|\chi|\right), \tag{13}$$

and

$$\underline{\gamma}_p\left(|\chi|\right) \geq \underline{c}_p\, |\chi|^a, \qquad \bar{\gamma}_p\left(|\chi|\right) \leq \bar{c}_p\, |\chi|^a. \tag{14}$$

where $\chi = \hat{x}_{k-N|k} - \bar{x}_{k-N}$, $0 \leq \underline{c}_p \leq \bar{c}_p$ and $a \in R_{\geq 2}$.

Given the updating scheme (8), inequality (13) verifies

$$|P_0^{-1}|\chi^a \leq \Gamma_{k-N}\left(\chi\right) \leq |P_\infty^{-1}|\chi^a. \tag{15}$$

Assumption 2.2 The stage cost $\ell\left(\cdot,\cdot\right) : \mathbb{R}^{n_w} \times \mathbb{R}^{n_v} \to \mathbb{R}$ is a continuous function bounded by $\underline{\gamma}_w, \underline{\gamma}_v, \bar{\gamma}_w, \bar{\gamma}_v \in \mathcal{K}_\infty$ such that $\forall w \in \mathcal{W}$ and $v \in \mathcal{V}$ the following inequalities are satisfied

$$\underline{\gamma}_w\left(|w|\right) + \underline{\gamma}_v\left(|v|\right) \leq \ell\left(w, v\right) \leq \bar{\gamma}_w\left(|w|\right) + \bar{\gamma}_v\left(|v|\right). \tag{16}$$

These assumptions allow us to built upper and lower bounds of the elements of Ψ. Then, functions γ_1 and γ_2 (from Definition 2.1) are use to measure the distance between the disturbance sequences $\boldsymbol{w_1}, \boldsymbol{w_2}, \boldsymbol{v_1}$ and $\boldsymbol{v_2}$. They are related with $\underline{\gamma}_w\left(|w|\right), \underline{\gamma}_v\left(|v|\right), \bar{\gamma}_w\left(|w|\right)$ and $\bar{\gamma}_v\left(|v|\right)$ through the following inequalities

$$\gamma_1\left(3\underline{\gamma}_w^{-1}\left(|w|\right)\right) \leq c_1|w|^{\alpha_1},\ \gamma_2\left(3\underline{\gamma}_v^{-1}\left(|w|\right)\right) \leq c_2|v|^{\alpha_2}, \tag{17}$$

for $c_{1,2}, \alpha_{1,2} \in \mathbb{R}_{\geq 0}$. Finally, an assumption about the structure and properties of function $\beta(r, s)$ is introduced

Assumption 2.3 The function $\beta(r, s) \in \mathcal{KL}$ and satisfies the following inequality

$$\beta(r,s) \le c_\beta r^p s^{-q}, \tag{18}$$

for some $c_\beta, p, q \in \mathbb{R}_{\ge 0}$ and $q \ge p$.

Inequalities (17) and (18) have been used in previous works [25], [26] to bound the estimator performance metric Ψ. Finally, the definition of a robust global asymptotic stability (RGAS) is introduced to study the stability of a nonlinear system subject to bounded disturbances.

DEFINITION 2.2 Consider the system (3) subject to disturbances $w \in W(w_{max})$ and $v \in V(v_{max})$ for $w_{max}, v_{max} \in \mathbb{R}_{\ge 0}$, with prior estimate $\bar{x}_0 \in X(e_{max})$. The MHE equation (5) with adaptive arrival cost (8) is robustly globally asymptotically stable (RGAS) if there exists functions $\Phi \in \mathcal{KL}$ and $\pi_w, \pi_v \in \mathcal{K}$ such that $\forall x_0 \in X$ and $\forall \bar{x}_0 \in X_0$ the following inequality is satisfied $\forall k \in \mathbb{Z}_{\ge 0}$

$$|x_k - \hat{x}_k| \le \Phi(|x_0 - \bar{x}_0|, k) + \pi_w\left(\|w\|_{[0,k-1]}\right) + \pi_v\left(\|v\|_{[0,k-1]}\right). \tag{19}$$

In the next section we will show that if system (3) is i-IOSS and Assumptions 2.1 to 2.3 are fulfilled, the MHE estimator (5) with adaptive arrival cost (8) is RGAS. Furthermore, if the process and measurement disturbances are convergent (i.e., $w \in C_w, v \in C_v$), then $\hat{x}_{k|k} \to x_k$ as $k \to \infty$.

3 Robust stability of moving horizon estimation under bounded disturbances

With all the elements introduced in the previous section, we are ready to derive the main result: the RGAS of the proposed moving horizon estimator with an estimation horizon $\mathbb{N}_{min}$ for nonlinear detectable systems under bounded disturbances. Furthermore, $\mathcal{KL}$ and $\mathcal{K}$ functions exist such that (19) is valid with Φ, π_w and π_v for all estimation horizon $N \ge \mathbb{N}_{min}$.

THEOREM 3.1 Given the i-IOSS nonlinear system (3) with a prior estimate $\bar{x}_0 \in X_0$ of its unknown initial condition x_0 and bounded disturbances $w \in W(w_{max})$, $v \in V(v_{max})$, Assumptions 2.1 to 2.3 are fulfilled, the prior weighting Γ_{k-N} is updated with algorithm (8) and the estimation horizon verifies

$$\mathbb{N}_{min} := \left\lceil \delta^{\frac{\zeta}{\eta}} e_{max}^{\frac{\zeta-1}{\eta}} \mathbb{C}_P^{\frac{\rho}{\eta}} \left(c_\beta 18^p + \lambda_{min}^{\alpha_1}\left(P_0^{-1}\right)\left(c_1 3^{\alpha_1} + c_2 3^{\alpha_2}\right) + c_\beta 2^p \right)^{\frac{1}{\eta}} \right\rceil, \tag{20}$$

then the resulting MHE estimator is stable.

Proof. See Appendix 7.

The proof of Theorem 3.1 is constructive and provides an estimate of the estimation horizon N required to guarantee the stability of the MHE estimator analyzed in this work. The estimates $\mathbb{N}_{min}$ and functions $\bar{\beta}$, φ_w and φ_v are conservative, since their derivation involved conservative estimates of noises, errors, stage and arrival costs. Functions $\bar{\beta}(r, s)$, $\varphi_w(r)$ and $\varphi_v(r)$ are given by

$$\bar{\beta}(|x_{k-N} - \bar{x}_{k-N}|, j) := \begin{cases} k_\beta |x_{k-N} - \bar{x}_{k-N}|^\zeta \, \mathbb{C}_P^\rho \left(c_\beta\, 18^p + \lambda_{min}^{\alpha_1}\left(P_0^{-1}\right)(c_1\, 3^{\alpha_1} + c_2\, 3^{\alpha_2}) + c_\beta\, 2^p \right), & j = 0 \\ \frac{|x_{k-N} - \bar{x}_{k-N}|^\zeta}{j^\eta} \, \mathbb{C}_P^\rho \left(c_\beta\, 18^p + \lambda_{min}^{\alpha_1}\left(P_0^{-1}\right)(c_1\, 3^{\alpha_1} + c_2\, 3^{\alpha_2}) + c_\beta\, 2^p \right), & j \in \mathbb{Z}_{[1,N]} \end{cases} \tag{21}$$

$$\varphi_w(\|w\|) := \frac{c_\beta\, 18^p}{|P_0^{-1}|} \bar{\gamma}_w^{\frac{p}{a}}(\|w\|) + c_2\, 3^{\alpha_2} \bar{\gamma}_w^{\alpha_2}(\|w\|) + \gamma_1(6\|w\|) + \gamma_1\left(6\underline{\gamma}_w^{-1}(3\bar{\gamma}_w(\|w\|))\right), \tag{22}$$

$$\varphi_v(\|v\|) := \frac{c_\beta\, 18^p}{|P_0^{-1}|} \bar{\gamma}_v^{\frac{p}{a}}(\|v\|) + c_1\, 3^{\alpha_1} \bar{\gamma}_v^{\alpha_1}(\|v\|) + \gamma_2(6\|v\|) + \gamma_2\left(6\underline{\gamma}_v^{-1}(3\bar{\gamma}_v(\|v\|))\right), \tag{23}$$

such that the estimation residual can be written as follows

$$|x_k - \hat{x}_{k|k}| \le \bar{\beta}(|x_{k-N} - \bar{x}_{k-N}|, k) + \varphi_w(\|w\|) + \varphi_v(\|v\|) \quad \forall k \in \mathbb{Z}_{[0,N]}. \tag{24}$$

To guarantee the validity of previous results on the entire time horizon we must determine if the resulting system is robust globally stable. Firstly, we determine the decreasing rate of the function $\bar{\beta}(r,s)$ $\mathbb{N}_{min}$ samplings time in the future. Adopting an estimator with a window length $N = \mathbb{N}_{min}$ such that

$$\bar{\beta}(\delta r, N) \le \left(\frac{\mathbb{N}_{min}}{N}\right)^\eta r, \tag{25}$$

the effects of the initial conditions x_0 will vanish with a decreasing rate δ. As $k \to \infty$, $\hat{x}_{k|k}$ will entry to the bounded set $X_I(w,v) \subseteq X$ defined by the noises of the system w and v as follows

$$X_I(w,v) := \left\{ |x_k - \hat{x}_{k|k}| \le \delta(1+\mu)\left(\varphi_w(\|w\|) + \varphi_v(\|v\|)\right) \right\}, \quad \mu \in \mathbb{R}_{>0}. \tag{26}$$

This set defines the minimum size region that can be achieved by removing the error in initial conditions ($|x_0 - \bar{x}_0| \le e_{max}$). For any MHE with adaptive arrival cost and window length $N = \mathbb{N}_{min}$, two situations can be considered

- The estimator has removed the effects of x_0 on $\hat{x}_{k|k}$ such that $|x_{k+j} - \hat{x}_{k+j|k+j}| \in X_I(w,v)$ for all $j \in \mathbb{Z}_{>0}$, and
- The estimator has not removed the effects of x_0 on $\hat{x}_{k|k}$ such that $|x_{k+j} - \hat{x}_{k+j|k+j}| \notin X_I(w,v)$ for some $j \in \mathbb{Z}_{>0}$,

Assuming the first situation, moving forward N samples ahead equation (24), the following inequalities hold

$$\begin{aligned} |x_{k+N} - \hat{x}_{k+N}| &\le \bar{\beta}(|x_k - \bar{x}_k|, N) + \varphi_w(\|w\|) + \varphi_v(\|v\|), \\ &\le \frac{|x_k - \bar{x}_k|}{\delta} + \varphi_w(\|w\|) + \varphi_v(\|v\|), \\ &\le \delta(1+\mu)\left(\varphi_w(\|w\|) + \varphi_v(\|v\|)\right). \end{aligned} \tag{27}$$

This equation implies the fact that $|x_{k+j} - \hat{x}_{k+j|k+j}| \in X_I(w,v)$ $\forall j \in \mathbb{Z}_{\ge 0}$. In the other case, when the estimation error is outside of $X_I(w,v)$, equations (24) and (25) are recalled again

sinc(i) Research Institute for Signals, Systems and Computational Intelligence (sinc.unl.edu.ar)
Nestor N. Deniz, M. Murillo, G. Sanchez & L. Giovanini; "Robust stability of moving horizon estimation for non-linear systems with bounded disturbances using adaptive arrival cost"
IET Control Theory \& Applications, Vol. 14, No. 18, pp. 2879-2888, 2020.

and the following inequalities hold

$$\begin{aligned}|x_{k+N} - \hat{x}_{k+N}| \leq & \frac{|x_k - \hat{x}_k|}{\delta}\left(\frac{\mathbb{N}_{min}}{N}\right)^{\eta} + \phi_w\left(\|\boldsymbol{w}\|\right) + \\ & \phi_v\left(\|\boldsymbol{v}\|\right), \\ \leq & |x_k - \bar{x}_k|\left(\frac{\mathbb{N}_{min}}{N}\right)^{\eta}\left(\frac{2+\mu}{\delta\left(1+\mu\right)}\right). \end{aligned} \tag{28}$$

If

$$\delta > \frac{2+\mu}{1+\mu}, \tag{29}$$

then

$$\forall N \geq \mathbb{N}_{min} : \theta := \left(\frac{\mathbb{N}_{min}}{N}\right)^{\eta}\left(\frac{2+\mu}{\delta\left(1+\mu\right)}\right) < 1. \tag{30}$$

Equation (28), under condition and (29), reveals a contractive behaviour of the estimation error with a contraction factor θ. Then, for some finite time $k^* \in \mathbb{Z}_{>\mathbb{N}_{min}}$ the estimation error will decrease until

$$|x_{k^*} - \hat{x}_{k^*|k^*}| \in \mathcal{X}_I\left(w, v\right). \tag{31}$$

In equivalent formulations, equations (27) and (28) put in evidence the existence of a positive invariant set and a Lyapunov like function for the proposed estimator. From equation (28), one can see that for the case that the estimation error belong to the set $\mathcal{X}_I\left(w,v\right)^C \cap \mathcal{X}$, ($\mathcal{X}_I\left(w,v\right)^C$ denotes the complement of the set $\mathcal{X}_I\left(w,v\right)$) the estimation error decreases in a factor of θ every $\mathbb{N}_{min}$ sampling time.

Taking into account the general case in which $|x_k - \hat{x}_{k|k}| \in \mathcal{X}_I\left(w,v\right)$ for $k \in \mathbb{Z}_{\geq \mathbb{N}_{min}}$, following the same procedure as in [25], we could define $i := \lfloor \frac{k}{N} \rfloor$ (where $\lfloor \cdot \rfloor$ denotes the floor function) and $j := k$ *mod* N, therefore $k = iN + j$. Combining equations (27) and (28) and taking into account that $|x_k - \hat{x}_{k|k}| \leq \delta e_{\max}$ for $k \in \mathbb{Z}_{[0,N-1]}$, the following inequalities are obtained

$$\begin{aligned}|x_k - \hat{x}_{k|k}| & \leq \max\{|x_j - \bar{x}_j|\,\theta^i, \delta\left(1+\mu\right)\left(\phi_w\left(\|\boldsymbol{w}\|\right) + \phi_v\left(\|\boldsymbol{v}\|\right)\right)\}, \\ & \leq \max\{\theta^i\left(\bar{\beta}\left(|x_0 - \bar{x}_0|, j\right) + \phi_w\left(\|\boldsymbol{w}\|\right) + \phi_v\left(\|\boldsymbol{v}\|\right)\right), \delta\left(1+\mu\right)\left(\phi_w\left(\|\boldsymbol{w}\|\right) + \phi_v\left(\|\boldsymbol{v}\|\right)\right)\}, \\ & \leq \theta^i\bar{\beta}\left(|x_0 - \bar{x}_0|, j\right) + 2\left(1+\mu\right)\left(\phi_w\left(\|\boldsymbol{w}\|\right) + \phi_v\left(\|\boldsymbol{v}\|\right)\right). \end{aligned} \tag{32}$$

Equation (32) can be rewritten as follows

$$|x_k - \hat{x}_{k|k}| \leq \theta^i\bar{\beta}\left(|x_0 - \bar{x}_0|,\ j\right) + 2\left(1+\mu\right)\left(\phi_w\left(\|\boldsymbol{w}\|_{[0,k-1]}\right) + \phi_v\left(\|\boldsymbol{v}\|_{[0,k-1]}\right)\right), \tag{33}$$

which is just equation (19) with

$$\Phi\left(|x_0 - \bar{x}_0|, k\right) = \theta^i\bar{\beta}\left(|x_0 - \bar{x}_0|,\ j\right), \tag{34}$$

$$\pi_w\left(\|\boldsymbol{w}\|_{[0,k-1]}\right) = 2\left(1+\mu\right)\phi_w\left(\|\boldsymbol{w}\|_{[0,k-1]}\right), \tag{35}$$

$$\pi_v\left(\|\boldsymbol{v}\|_{[0,k-1]}\right) = 2\left(1+\mu\right)\phi_v\left(\|\boldsymbol{v}\|_{[0,k-1]}\right). \tag{36}$$

This equations imply the fact that the proposed estimator is *RGAS*.

REMARK 3.1 Functions ϕ_w and ϕ_v in equations (22) and (23), and hence π_w and π_v in equations (35) and (36), do not depend on the estimation horizon N which means that the moving horizon estimator with adaptive arrival cost is RGAS with uniform gains given by (22) and (23).

This fact implies that the size of the invariant set $\mathcal{X}_I(w,v)$ is independent of the horizon length N and it only depends on the a priori bounds for the maximum disturbances (see Example 2). Moreover, starting with a poor estimate of initial condition $\bar{x}_0$, estimation error will decrease thanks to the arrival–cost updating mechanism, which improves the prior estimate $\bar{x}_0$. When the knowledge about the state at the beginning of the estimation window is accurate (i.e., function (34) is near zero), one can considerate that the estimation error has entered in the invariant space $\mathcal{X}_I\left(w,v\right)$, reaching its minimum value determined by functions (35) and (36).

REMARK 3.2 Equation (33) also holds for $\|\boldsymbol{w}\|_{[k-N,k-1]}$ and $\|\boldsymbol{v}\|_{[k-N,k]}$ instead of $\|\boldsymbol{w}\|$ and $\|\boldsymbol{v}\|^{\dagger}$.

This fact is consequence of the exponential averaging introduced by variable forgetting factor α_k. This fact improves the robustness and performance of the estimator by adapting the relative significance of the arrival cost Γ_{k-N} on the estimate $\hat{x}_{k|k}$. The effect is to generate more accurate estimates of ϕ_w and ϕ_v, since they depend on the information available within the estimation window.

Finally, the convergence of the estimation error $|x_k - \hat{x}_{k|k}| \to 0$ is shown for $w_k \in \mathcal{C}_w$ and $v_k \in \mathcal{C}_v$. Since $\Phi\left(|x_0 - \bar{x}_0|,\ k\right) \in \mathcal{KL}$ by construction and sequences $\boldsymbol{w}$ and $\boldsymbol{v}$ are convergent, the right hand side of equation (33) can be rewritten taking into account only the disturbances within the estimation window, i.e., $\|\boldsymbol{w}\|_{[k-N,k-1]}$ and $\|\boldsymbol{v}\|_{[k-N,k-1]}$. Since $\lim_{k\to\infty} w_k = 0$ and $\lim_{k\to\infty} v_k = 0$, one can choose some k_1 large enough such that $\max\left\{w_{\max}, v_{\max}\right\} \leq \min\left\{\pi_w^{-1}\left(\frac{\varepsilon}{3}\right), \pi_v^{-1}\left(\frac{\varepsilon}{3}\right)\right\}$. Since $\Phi\left(|x_0 - \bar{x}_0|, k\right) \in \mathcal{KL}$, there exist some k_2 such that $\Phi\left(\cdot, k_2\right) \leq \frac{\varepsilon}{3}$. Taking $k \geq \max\{k_1, k_2\}$, $|x_k - \hat{x}_{k|k}| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$. Since ε is arbitrary, $|x_k - \hat{x}_{k|k}| \to 0$ when $k \to \infty$ (see Figure 5 from example 2). Note that to prove convergence to the true state, one has to take into account only the disturbances acting on the horizon of the estimator. On the other hand, to prove *RGAS* is necessary to take into account all the history of disturbances.

For the case of non-vanishing disturbances, whenever the horizon length is chosen as $N \geq \mathbb{N}_{min}$, the minimum estimation error is determined by functions (35) and (36). However, when $w_k \in \mathcal{C}_w$ and $v_k \in \mathcal{C}_v$, we have shown that $|x_k - \hat{x}_{k|k}| \to 0$. This seems to be an unrealistic situation. However it entails that in a noiseless (ideal) scenario, the algorithm does not introduce artificial noise, and the estimation error will be upper bounded only by equation (34), which behaves asymptotically stable to zero whenever $N \geq \mathbb{N}_{min}$.

Note that $\mathbb{N}_{min}$, which guarantees the *RGAS* of the system, depends on e_{max}, the class of disturbances considered (upper bounds of w and v), P_0 and the bounds of the stage cost, while it is independent of $\|\boldsymbol{w}\|$, $\|\boldsymbol{v}\|$. Therefore, the same $\mathbb{N}_{min}$ ensures *RGAS* for all bounded disturbances and e_{max}. This implies that we can prove the *RGAS* property for full information estimator with least–square objective function.

†*It can be done omitting last step in equation* (42).

sinc(*i*) Research Institute for Signals, Systems and Computational Intelligence (sinc.unl.edu.ar)
Nestor N. Deniz, M. Murillo, G. Sanchez & L. Giovanini; "Robust stability of moving horizon estimation for non-linear systems with bounded disturbances using adaptive arrival cost"
IET Control Theory \& Applications, Vol. 14, No. 18, pp. 2879-2888, 2020.

4 Examples

The following examples will be used to illustrate the results presented in the previous sections and compare the performance of the estimators. The examples considered in this work are taken from [25] for a direct comparison of the results.

4.1 Example 1

The first example considers the system

$$x(k+1) = \begin{bmatrix} 0.8x_1(k) + 0.2x_2(k) + 0.5w(k) \\ -0.3x_1(k) + 0.5\cos(x_2(k)) \end{bmatrix} \qquad (37)$$
$$y(t) = x_2(t) + v(t)$$

The stage cost is chosen as $\ell(w, v) = 10w^2 + 10v^2$ and the horizon length is $N = 10$ for all estimators. The arrival cost is chosen as $\Gamma(\chi) = 0.1(\chi - \hat{x}(t|t))^T(\chi - \hat{x}(t|t))$ for the MAX estimator ([25]) and $\Gamma_t(\chi) = (\chi - \hat{x}(t|t))^T\Pi_k^{-1}(\chi - \hat{x}(t|t))$ for the ADAP estimator respectively, where $\Pi_0 = 10I_2$ and Π_k is computed using equations (8) with $\sigma = 0.2$ and $c = 1e6$. The MAX estimator uses $\delta = 1$, $\delta_1 = \kappa^N$ with $\kappa = 0.89^2$ and $\delta_2 = 1/N$ (see equation (3) of [25]). The full information estimator (FIE MAX, see [26]) is configured with the same parameter used by [25], arrival cost Γ_0 and $\delta = 1$, $\delta_1 = \kappa^k$ and $\delta_2 = 1/k$.

Table 1 Example 1 averaged MSE.

	FIEMAX	ADAP	MAX	EKF
x_1	0.02040	0.02176	0.02206	0.02296
x_2	0.00135	0.00151	0.00156	0.00154

Table 1 shows the mean square estimation error of each estimator averaged over 300 trials. It can be seen that the proposed estimator average mean square estimation error is smaller than MAX ones and closer to FIEMAX. The main performance difference between ADAP and FIEMAX estimators is the inclusion of the $\max$ term in the last one, which allows to follow the sudden changes in the signals (see Figure 1). The EKF provides the worst performance of the estimators analyzed in this example. The FIE estimator is the ones that uses the bigger amount of information (all the samples available). On the other hand, the EKF uses the smaller amount of information (the actual sample). This is the main reason why the FIEMAX provides the best performance and EKF provides the worst performance of the estimators considered in this example. The ADAP and MAX estimators use the same amount of information (N samples) but they process the information of previous samples in different way. This is the main reason for the difference of performance between these estimators.

Figure 1 shows simulation results for initial condition $x_0 = [0.5, 0]^T$ and prior estimate $\bar{x}_0 = [0, 0]^T$. The process and measurement disturbances w and v are sampled from an uniform distribution over the intervals $[-0.3, 0.3]$ and $[-0.2, 0.2]$, respectively. This figure shows that the estimators that use the $\max$ term (FIEMAX and ADAP) are able of following the sudden changes, however in the remaining of the signal the MAX estimator is moving away of the FIEMAX while ADAP remains closer it. This behaviour is due to the presence of the $\max$ in the performance index of the MAX estimator, while the ADAP estimator smooths the estimates $\hat{x}_{k|k}$. It is interesting to point out that the three receding horizon estimators (FIEMAX, MAX and ADAP) have the same behaviour during the first part of transient behaviour ($k < 5$). During the second part of transient behaviour ($5 < k < 10$) the FIEMAX and MAX estimators still have the same behaviour while the ADAP move away. Finally, after the transient behaviour ($k > 10$), the three estimators have a different behaviours. These behaviours are explained by the structure of the performance index and the amount of information used by each estimator, as it was explained in the previous paragraph.

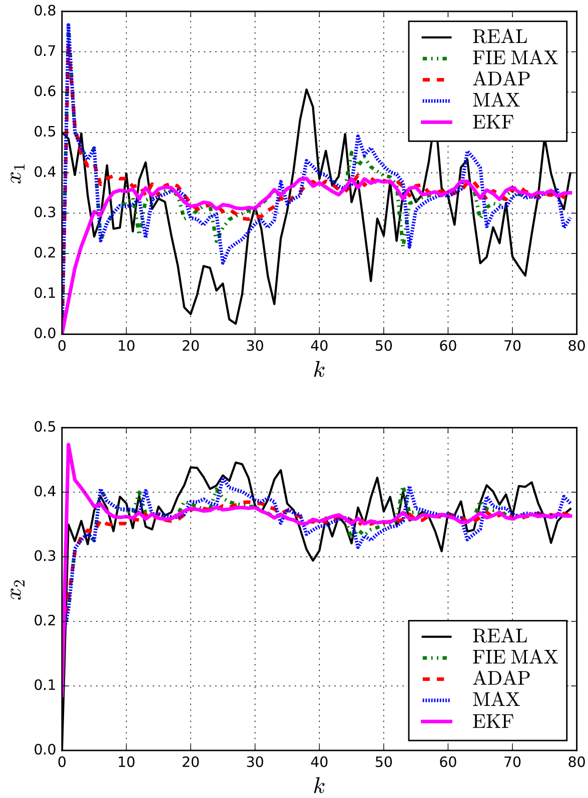

Fig. 1: Comparison between different estimators and system states.

4.1.1 MHE in the presence of variable measurement noise:
Now the MHE estimator is evaluated in the presence of time-varying measurement noise. The variance of the measurement noise is changed from 0.2 to 1.0 between times 20 and 40, then it returns to 0.2.

Table 2 Example 1 average MSE with variable measurement noise.

	FIEMAX	ADAP	MAX
x_1	0.02060	0.02068	0.03067
x_2	0.00168	0.00290	0.00335

Table 2 shows the average mean square error in the presence of variable measurement noise. In this case we can see that the mean square error of the ADAP estimator is marginally affected (less than 2%) by the changes of the measurement noise, while the mean square error of the MAX estimator increases significantly (more than 40%). The behaviour of ADAP estimator performance is due to the adaptation capabilities of the prior weighting updating mechanism, which is able of tracking the changes of noises by adapting the arrival cost $\Gamma_{k-N|k}$. In the case of MAX estimator, the arrival cost is updated using information of the system model and noises [25], therefore any change in the operational conditions is not incorporated into the arrival cost, affecting the remaining estimates of the window. In the case of the FIEMAX estimator, its performance is not affected (less than 2%) because the FIEMAX uses all the information available until the current sample.

Figure 2 shows the evolution of the trace of P_{k-N}^{-1} used in the prior weight of ADAP estimator in both cases, fix and variable measurement noise. It can be seen that the trace of both matrices grow in similar way, however when the measurement noise changes its variance from 0.2 to 1.0 the trace of P_{k-N}^{-1} increases its value (from 12.5 to 22.5) and them both traces have the same behaviour again. This change in the arrival cost weight helps ADAP to acquire information that allows it to keep the performance (from $MSE = 0.02181$ to $MSE = 0.02088$) while the MSE of MAX increases a 40%

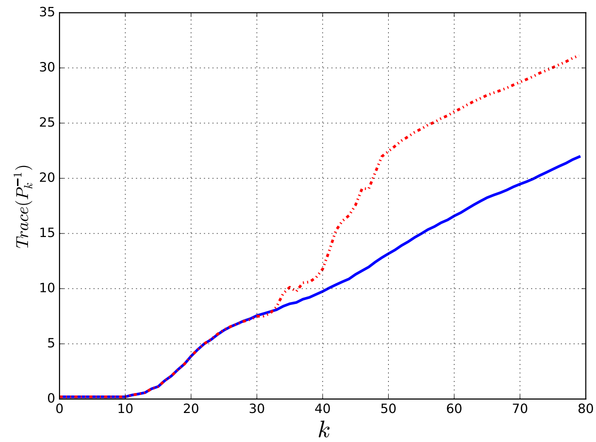

Fig. 2: Comparison of the evolution of $trace(P_{k-N}^{-1})$ used by ADAP estimator for time–varying (red dash dotted) and constant (blue dashed) measurement noise parameters.

Table 3 Example 2 averaged MSE over 300 trials and different horizon size.

	N=2		N=5		N=10		
	ADAP	MAX	ADAP	MAX	ADAP	MAX	FIE
x_1	0.18808	0.58652	0.03367	0.04615	0.00171	0.00772	0.00024
x_2	0.23037	0.66768	0.04074	0.05077	0.00285	0.00951	0.00120

(from $MSE = 0.02211$ to $MSE = 0.03085$). The trace of P_{k-N}^{-1} keeps changing until $k = 50$ since the last measurement affected by the noise with variance 1.0 is abandoning the estimation window (remember $N = 10$).

4.2 Example 2

As a second example, we consider a second order gas-phase irreversible reaction of the form $2A \rightarrow B$. This example has been considered in the context of moving horizon estimation in [33], [26] and [25]. Assuming an isothermal reaction and that the ideal gas law holds, the system dynamics

$$\begin{aligned} \dot{x}(t) &= \begin{bmatrix} -2\kappa x_1^2(t) + w(t) \\ \kappa x_1^2(t) \end{bmatrix}, \\ h(x) &= x_1(t) + x_2(t) + v(t), \end{aligned} \tag{38}$$

where $x = [x_1, x_2]$, x_1 is the partial pressure of the reactant A, x_2 is the partial pressure of the product B, and $\kappa = 0.16$ is the reaction rate constant. The measured output of the system is the total pressure. The system is affected with additive process and measurement noise w and v drawn from normal distributions with zero mean and covariance $Q_w = 0.001^2 I_2$ and $R_v = 0.1^2$, respectively. The stage cost and prior weighting are chosen as $\ell(w,v) = w^T Q_w^{-1} w + R_v^{-1} v^2$ and $\Gamma_t(\chi) = (\chi - \hat{x}(k|k))^T \Pi_k^{-1}(\chi - \hat{x}(k|k)$ with $\Pi_0 = (1/36) I_2$, where Π_k is determined by an extended Kalman filtering recursion in the case of the MAX estimator and the adaptive method in the case of the ADAP estimator with $\sigma = 0.1$ and $c = 1e6$. For the MAX estimator we use $\delta_1 = 1/N$, $\delta_2 = 1$ and $\delta = 0$. In the case of the ADAP estimator, the stage cost weight matrices are chosen as $Q_w = 0.001 I_2$ and $R_v = 0.1$. We use a multiple shooting strategy with a sampling time of $\Delta = 0.1\,[Seg]$ and we add the restrictions $x_1 \geq 0$ and $x_2 \geq 0$.

Table 3 shows the values of the mean squared error computed from the time 10 (in order to neglect the initial transient error) up to the simulation end time and averaged over 300 trials for horizon sizes of $N = 2$, $N = 5$ and $N = 10$.

The results show that the performance of ADAP estimator is superior to the one of MAX estimator for any horizon (two to three times smaller), showing a decrement of the absolute value of MSE with the size of the estimation window for both estimator. The FIE provides the best performance at the price of its computational burden. Since the FIE has into account all measurements, as time increase the problem become intractable from a practical point of view. However, it is useful for benchmark purposes. In this example we can see again the effect of the arrival cost on the performance of MHE estimator, playing a central role on the behaviour of the estimators. This effect is stronger for shorter horizons, since the significance of the arrival cost on the cost function decreases with the size of the estimation window. The ADAP estimator always provides a better performance than the MAX, providing performance improvements ranging from 277% (for $N = 10$) to 588% (for $N = 5$).

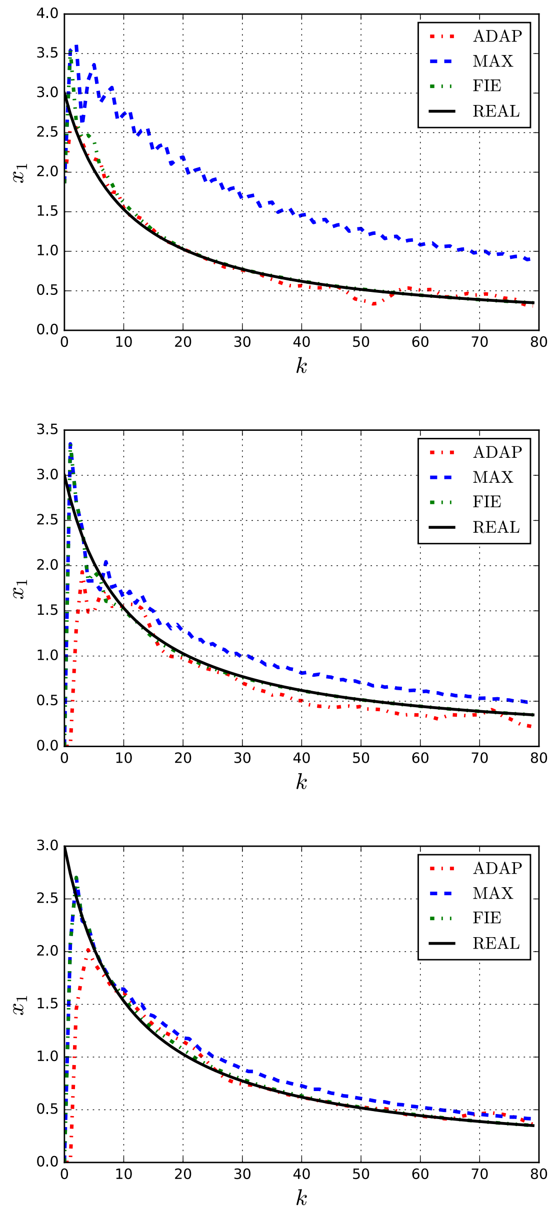

Fig. 3: Comparison of x_1 between ADAP, MAX and FIEMAX estimators for different horizon length (top $N = 2$, middle $N = 5$ and bottom $N = 10$).

Figures 3 and 4 show simulation results for $x_0 = [3, 1]^T$, $\bar{x}_0 = [0.1, 4.5]^T$ and estimation horizons of sizes $N = 2, 5$ and 10, along with results for a full information estimator using the same parameters. These figures show that the behaviour of ADAP estimator hardly change with horizon length (only the startup behaviours show significant differences), providing estimates $\hat{x}_{k|k}$ that converge to a

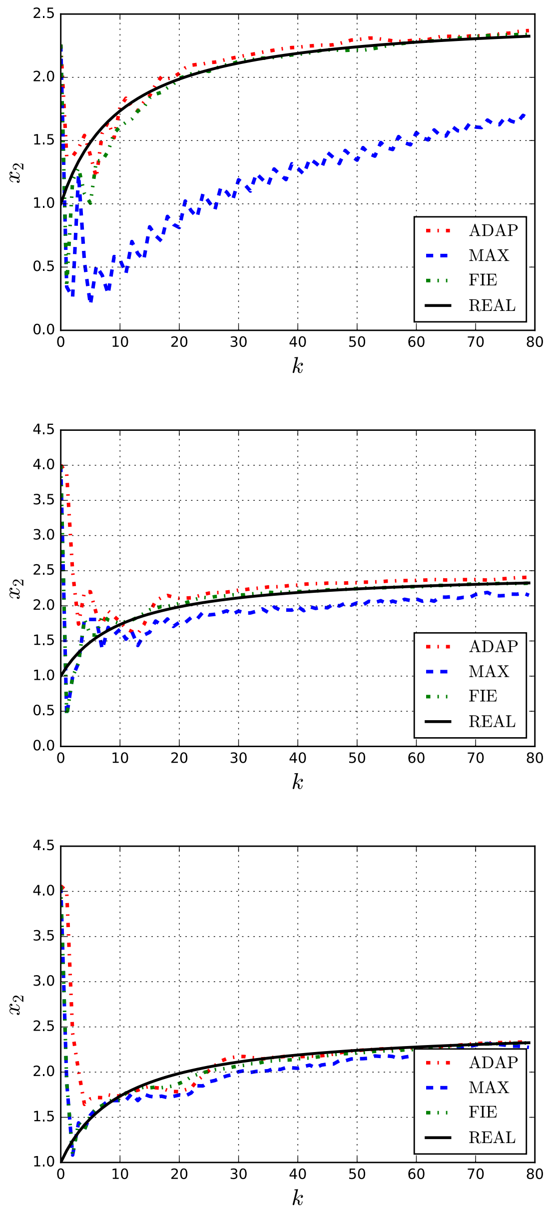

Fig. 4: Comparison of x_2 between ADAP, MAX and FIEMAX estimators for different horizon length (top $N = 2$, middle $N = 5$ and bottom $N = 10$).

neighbourhood of the true states x_k, whose size depends on the variance of the noises the initial value of the arrival cost weight (P_0) and the stage costs (see equations (22) and (23)). This result is consistent with the independence of observer gains π_w and π_v from the size of estimation window N (see Remark 3.1). On the other hand, the behaviour of the MAX estimator changes significantly with the size of the estimation window. In all the cases it provides estimates $\hat{x}_{k|k}$ that do not converge to the true states x_k, showing an offset that depends on the size of the estimation window. This result is consistent with the dependence of observer gains π_w and π_v with the estimation window size N (see Remark 16 of [25]). In addition, the MAX estimator also exhibits the cycling effect caused by the use of the filtered estimate to update $\bar{x}_{k-N}$ [29].

These figures also show that the main difference in the performance between ADAP and FIE is due to the transient phase ($0 < k < 10$), while during the steady state phase ($k > 10$) both estimators have a similar behaviour. On the other hand, MAX and FIE estimators show different behaviours along all simulation.

Finally, Figure 5 shows through simulations the convergence of $\hat{x}_{k|k}$ to x_k when disturbances are convergent. In these simulations the only source of error is $\bar{x}_0 \neq x_0$, therefore sequences of noises are convergent $\boldsymbol{w} \in \mathcal{C}_w$ and $\boldsymbol{v} \in \mathcal{C}_v$. In this figure we can see that independently of the initial condition $\bar{x}_0$, $\hat{x}_{k|k}$ converges to x_k.

aaa

5 Conclusions

In this paper we established robust global asymptotic stability for moving horizon estimator with a least-square type cost function for nonlinear detectable (i-IOSS) systems in presence of bounded disturbances. It was also shown that the estimation error converges to zero in case that disturbances converge to zero. This was done for an estimator which uses a least-square type cost function whose arrival cost is updated using adaptive estimation methods. An advantage of this updating mechanism is that the required conditions on prior weighting are such that it can be chosen off-line. Furthermore, it introduces a feedback mechanism between the arrival cost weight and the estimation errors that automatically controls the amount of information used to compute it, which allows to shorten the estimation horizon.

The standard least-square type cost function is typically used in practical applications and RGAS has been proved in [25]. However, for this formulation, the disturbances gains depend on the estimation horizon. Hence, this result does not allow to establish robust global asymptotic stability for a full information estimator. We showed that changing the updating mechanism of arrival cost weight the disturbances gains becomes uniform, allowing to extend the stability analysis to full information estimators with least-square type cost functions.

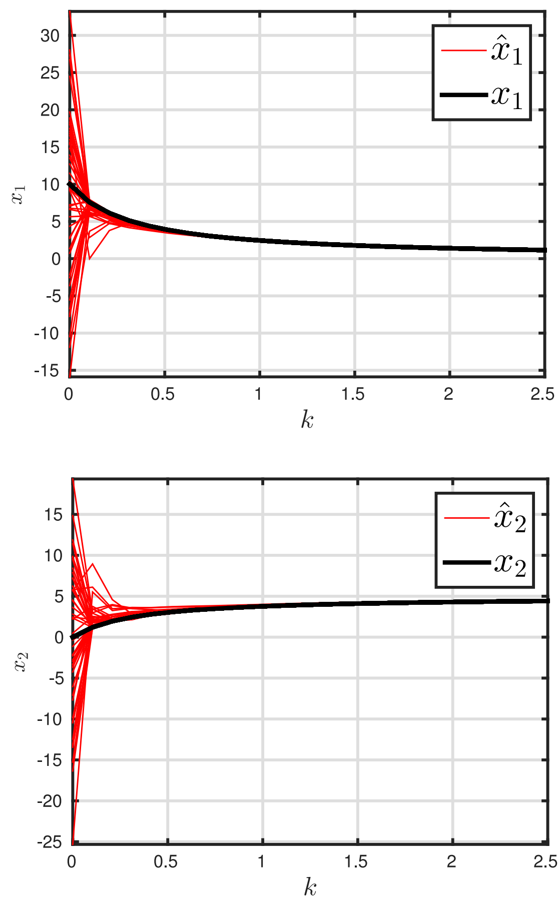

Fig. 5: Convergence of the estimation error for the case of convergent disturbances.

Acknowledgment

The authors wish to thank the Consejo Nacional de Investigaciones Cientificas y Tecnicas (CONICET) from Argentina, for their support.

6 References

1 Andrew H Jazwinski. *Stochastic processes and filtering theory*. Courier Corporation, 2007.
2 John L Crassidis and John L Junkins. *Optimal estimation of dynamic systems*. Chapman and Hall/CRC, 2004.
3 Huaizhong Li and Minyue Fu. A linear matrix inequality approach to robust h/sub/spl infin//filtering. *IEEE Transactions on Signal Processing*, 45(9):2338–2350, 1997.
4 Ali H Sayed. A framework for state-space estimation with uncertain models. *IEEE Transactions on Automatic Control*, 46(7):998–1013, 2001.
5 Franco Blanchini and Stefano Miani. Set-theoretic methods in control. systems & control: Foundations & applications. *Birkhäuser. Boston, MA*, 2008.
6 Laurent El Ghaoui and Hervé Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on matrix analysis and applications*, 18(4):1035–1064, 1997.
7 K Hu and J Yuan. Improved robust h inf. filtering for uncertain discrete-time switched systems. *IET Control Theory & Applications*, 3(3):315–324, 2009.
8 Xing Zhu, Yeng Chai Soh, and Lihua Xie. Design and analysis of discrete-time robust kalman filters. *Automatica*, 38(6):1069–1077, 2002.
9 A Jazwinski. Limited memory optimal filtering. *IEEE Transactions on Automatic Control*, 13(5):558–563, 1968.
10 Fred C Schweppe. *Uncertain dynamic systems*. Prentice Hall, 1973.
11 Christopher V Rao, James B Rawlings, and Jay H Lee. Constrained linear state estimationâĂŤa moving horizon approach. *Automatica*, 37(10):1619–1628, 2001.
12 Christopher V Rao, James B Rawlings, and David Q Mayne. Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE transactions on automatic control*, 48(2):246–258, 2003.
13 Angelo Alessandri, Marco Baglietto, and Giorgio Battistelli. Robust receding-horizon state estimation for uncertain discrete-time linear systems. *Systems & Control Letters*, 54(7):627–643, 2005.
14 Angelo Alessandri, Marco Baglietto, and Giorgio Battistelli. Moving-horizon state estimation for nonlinear discrete-time systems: New stability results and approximation schemes. *Automatica*, 44(7):1753–1765, 2008.
15 Angelo Alessandri, Marco Baglietto, and Giorgio Battistelli. Min-max moving-horizon estimation for uncertain discrete-time linear systems. *SIAM Journal on Control and Optimization*, 50(3):1439–1465, 2012.
16 J Garcia-Tirado, H Botero, and F Angulo. A new approach to state estimation for uncertain linear systems in a moving horizon estimation setting. *International Journal of Automation and Computing*, 13(6):653–664, 2016.
17 Hossein Sartipizadeh and Tyrone L Vincent. Computationally tractable robust moving horizon estimation using an approximate convex hull. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pages 3757–3762. IEEE, 2016.
18 G Sánchez, M Murillo, and L Giovanini. Adaptive arrival cost update for improving moving horizon estimation performance. *ISA transactions*, 68:54–62, 2017.
19 James B Rawlings and David Q Mayne. Model predictive control: Theory and design. 2009.
20 James B Rawlings and Luo Ji. Optimization-based state estimation: Current status and some new results. *Journal of Process Control*, 22(8):1439–1444, 2012.
21 Douglas A Allan and James B Rawlings. Moving horizon estimation. In *Handbook of Model Predictive Control*, pages 99–124. Springer, 2019.
22 Angelo Alessandri, Marco Baglietto, Giorgio Battistelli, and Victor Zavala. Advances in moving horizon estimation for nonlinear systems. In *Decision and Control (CDC), 2010 49th IEEE Conference on*, pages 5681–5688. IEEE, 2010.
23 Eduardo D Sontag and Yuan Wang. Output-to-state stability and detectability of nonlinear systems. *Systems & Control Letters*, 29(5):279–290, 1997.
24 Wuhua Hu, Lihua Xie, and Keyou You. Optimization-based state estimation under bounded disturbances. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 6597–6602. IEEE, 2015.
25 Matthias A Müller. Nonlinear moving horizon estimation in the presence of bounded disturbances. *Automatica*, 79:306–314, 2017.
26 Luo Ji, James B Rawlings, Wuhua Hu, Andrew Wynn, and Moritz Diehl. Robust stability of moving horizon estimation under bounded disturbances. *IEEE Transactions on Automatic Control*, 61(11):3509–3514, 2016.
27 Wuhua Hu. Robust stability of optimization-based state estimation. *arXiv preprint arXiv:1702.01903*, 2017.
28 Sven Knüfer and Matthias A Müller. Robust global exponential stability for moving horizon estimation. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 3477–3482, 2018.
29 Peter Klaus Findeisen. *Moving horizon state estimation of discrete time systems*. PhD thesis, University of Wisconsin–Madison, 1997.
30 Eduardo D Sontag. Some connections between stabilization and factorization. In *Decision and Control, 1989., Proceedings of the 28th IEEE Conference on*, pages 990–995. IEEE, 1989.
31 Z-P Jiang, Andrew R Teel, and Laurent Praly. Small-gain theorem for iss systems and applications. *Mathematics of Control, Signals and Systems*, 7(2):95–120, 1994.
32 Eduardo D Sontag. Input to state stability: Basic concepts and results. In *Nonlinear and optimal control theory*, pages 163–220. Springer, 2008.
33 Eric L Haseltine and James B Rawlings. Critical evaluation of extended kalman filtering and moving-horizon estimation. *Industrial & engineering chemistry research*, 44(8):2451–2460, 2005.

7 Appendix: Proof Theorem 1

The following optimal cost of problem (5) can be given by

$$\Psi_N^* = \Gamma_{k-N}\left(\chi^*\right) + \sum_{j=k-N}^{k} \ell\left(\hat{w}_{j|k}^*, \hat{v}_{j|k}^*\right), \tag{39}$$

which is bounded (Assumptions 2.1 and 2.2) $\forall\, |\hat{w}_{j|k}|$ and $\forall\, |\hat{v}_{j|k}|$ for all $j \in \mathbb{Z}_{[k-N,k]}$ by

$$\Psi_N^* \leq \bar{\gamma}_p\left(\chi^*\right) + N\bar{\gamma}_w\left(|\hat{w}_{j|k}^*|\right) + N\bar{\gamma}_v\left(|\hat{v}_{j|k}^*|\right), \tag{40}$$

$$\Psi_N^* \geq \underline{\gamma}_p\left(\chi^*\right) + N\underline{\gamma}_w\left(|\hat{w}_{j|k}^*|\right) + N\underline{\gamma}_v\left(|\hat{v}_{j|k}^*|\right). \tag{41}$$

Due to optimality, the following inequalities hold $\forall k \in [k-N, k-1]$

$$\begin{aligned}\Psi\left(\hat{x}_{k-N|k}^*, \hat{\boldsymbol{w}}^*\right) &\leq \Psi\left(x_{k-N}, \boldsymbol{w}\right), \\ &\leq \bar{\gamma}_p\left(|x_{k-N} - \bar{x}_{k-N}|\right) + N\bar{\gamma}_w\left(\|\boldsymbol{w}\|\right) + N\bar{\gamma}_v\left(\|\boldsymbol{v}\|\right),\end{aligned} \tag{42}$$

then, taking into account the lower and upper bounds we have

$$|\chi| \leq \underline{\gamma}_p^{-1}\left(\bar{\gamma}_p\left(|x_{k-N} - \bar{x}_{k-N}|\right) + N\bar{\gamma}_w\left(\|\boldsymbol{w}\|\right) + N\bar{\gamma}_v\left(\|\boldsymbol{v}\|\right)\right).$$

By mean of Assumptions 2.1 and 2.2 the last inequality can be written as follows

$$\begin{aligned}|\chi| &\leq \underline{\gamma}_p^{-1}\left(3\,\bar{\gamma}_p\left(|x_{k-N} - \bar{x}_{k-N}|\right)\right) + \underline{\gamma}_p^{-1}\left(3N\bar{\gamma}_w\left(\|\boldsymbol{w}\|\right)\right) + \underline{\gamma}_p^{-1}\left(3N\bar{\gamma}_v\left(\|\boldsymbol{v}\|\right)\right), \\ &\leq \frac{3^{\frac{1}{a}}}{|P_0^{-1}|}\left(|P_\infty^{-1}|^{\frac{1}{a}}|x_{k-N} - \bar{x}_{k-N}| + N^{\frac{1}{a}}\bar{\gamma}_w^{\frac{1}{a}}\left(\|\boldsymbol{w}\|\right) + N^{\frac{1}{a}}\bar{\gamma}_v^{\frac{1}{a}}\left(\|\boldsymbol{v}\|\right)\right).\end{aligned} \tag{43}$$

Analogously, bounds for $|\hat{v}_{j|k}|$ and $|\hat{w}_{j|k}|$ can be found

$$
\begin{aligned}
|\hat{w}_{j|k}| &\leq \underline{\gamma}_w^{-1}\left(\frac{3}{N}\bar{\gamma}_p\left(|x_{k-N}-\bar{x}_{k-N}|\right)\right)+\underline{\gamma}_w^{-1}\left(3\,\bar{\gamma}_w\left(\|\boldsymbol{w}\|\right)\right)+\underline{\gamma}_w^{-1}\left(3\,\bar{\gamma}_v\left(\|\boldsymbol{v}\|\right)\right),\\
|\hat{v}_{j|k}| &\leq \underline{\gamma}_v^{-1}\left(\frac{3}{N}\,\bar{\gamma}_p\left(|x_{k-N}-\bar{x}_{k-N}|\right)\right)+\underline{\gamma}_v^{-1}\left(3\,\bar{\gamma}_w\left(\|\boldsymbol{w}\|\right)\right)+\underline{\gamma}_v^{-1}\left(3\,\bar{\gamma}_v\left(\|\boldsymbol{v}\|\right)\right).
\end{aligned}
\tag{44}
$$

Next, let us consider some samples $k\in\mathbb{Z}_{\geq N}$ and assume that system (3) is ***i-IOSS*** with $z_1=x_{k-N}$, $z_2=\hat{x}_{k-N|k}$, $w_1=\{w_j\}$, $w_2=\{\hat{w}_{j|k}\}$, $v_1=\{v_j\}$ and $v_2=\{\hat{v}_{j|k}\}$ for all $j\in\mathbb{Z}_{[k-N,k-1]}$. Since $x(k)=x\left(N,z_1,\boldsymbol{w_1}\right)$, $\hat{x}(k)=\hat{x}_{k|k}=x\left(N,z_2,\boldsymbol{w_2}\right)$ we obtain

$$
|x_k-\hat{x}_{k|k}|\leq\beta\left(|x_{k-N}-\hat{x}_{k-N|k}|,\ N\right)+\gamma_1\left(\|\boldsymbol{w}_j-\hat{\boldsymbol{w}}_{j|k}\|\right)+\gamma_2\left(\|\boldsymbol{v}_j-\hat{\boldsymbol{v}}_{j|k}\|\right).
\tag{45}
$$

In order to get a finite upper bound for the estimation error, the three terms in the right hand side of equation (45) must be upper bounded. The first term can be written

$$
\begin{aligned}
\beta\left(|x_{k-N}-\hat{x}_{k-N|k}|,\ N\right) &\leq \beta\left(2\,|x_{k-N}-\bar{x}_{k-N}|,\ N\right)+\beta\left(2\,|\chi|,\ N\right),\\
&\leq \beta\left(2\,|x_{k-N}-\bar{x}_{k-N}|,\ N\right)+\beta\left(\frac{6\,3^{\frac{1}{a}}|P_\infty^{-1}|^{\frac{1}{a}}}{|P_0^{-1}|}\,|x_{k-N}-\bar{x}_{k-N}|,\ N\right)+\\
&\quad\beta\left(\frac{6\,3^{\frac{1}{a}}\ N^{\frac{1}{a}}}{|P_0^{-1}|}\,\bar{\gamma}_w^{\frac{1}{a}}\left(\|\boldsymbol{w}\|\right),\ N\right)+\beta\left(\frac{6\,3^{\frac{1}{a}}\ N^{\frac{1}{a}}}{|P_0^{-1}|}\,\bar{\gamma}_v^{\frac{1}{a}}\left(\|\boldsymbol{v}\|\right),N\right).
\end{aligned}
\tag{46}
$$

Using Assumptions 2.1 and 2.3, function $\beta(\cdot)$ is bounded by

$$
\begin{aligned}
\beta\left(|x_{k-N}-\hat{x}_{k-N|k}|,N\right) &\leq \frac{c_\beta\,2^p}{N^q}\,|x_{k-N}-\bar{x}_{k-N}|^p+\left(\frac{|P_\infty^{-1}|}{|P_0^{-1}|}\right)^p\frac{c_\beta\,6^p\,3^{\frac{p}{a}}}{N^q}\,|x_{k-N}-\bar{x}_{k-N}|^p+\\
&\quad\frac{c_\beta\,6^p\,3^{\frac{p}{a}}\,N^{\frac{p}{a}-q}}{|P_0^{-1}|}\,\left(\bar{\gamma}_w^{\frac{p}{a}}\left(\|\boldsymbol{w}\|\right)+\bar{\gamma}_v^{\frac{p}{a}}\left(\|\boldsymbol{v}\|\right)\right).
\end{aligned}
\tag{47}
$$

Taking into account that P_k^{-1} is a symmetric positive definite matrix $\forall k\in\mathbb{Z}_{[0,\infty)}$, then $|P_k^{-1}|\leq\lambda_{\max}\left(P_k^{-1}\right)$, where $\lambda_{\max}\left(P_k^{-1}\right)$ denotes the maximal eigenvalue of matrix P_k^{-1}. Denoting $\lambda_{\min}\left(P_k^{-1}\right)$ as the minimal eigenvalue of matrix P_k^{-1} and taking into account that $|P_k^{-1}|\leq|P_{k+1}^{-1}|$, the maximum conditioning number of matrix P_k^{-1} can be defined as $\mathbb{C}_P:=\lambda_{\max}\left(P_\infty^{-1}\right)/\lambda_{\min}\left(P_0^{-1}\right)$, then

$$
\beta\left(|x_{k-N}-\hat{x}_{k-N|k}|,\ N\right)\leq\frac{c_\beta}{N^q}\left(2^p+\mathbb{C}_P^p18^p\right)|x_{k-N}-\bar{x}_{k-N}|^p+\frac{c_\beta\,18^p}{|P_0^{-1}|}\left(\bar{\gamma}_w^{\frac{p}{a}}\left(\|\boldsymbol{w}\|\right)+\bar{\gamma}_v^{\frac{p}{a}}\left(\|\boldsymbol{v}\|\right)\right).
\tag{48}
$$

The first term in the right side of this equation is bounded due to the assumption that $|x_{k-N}-\bar{x}_{k-N}|\in\mathcal{X}_0\left(e_{\max}\right)$, while the second term are finite constants. To guarantee the validity of previous results on the entire time horizon we extend the definition of $\beta\left(r,s\right)$ to $N=0$. It is sufficient to define $\bar{\beta}\left(r,0\right)$ as follows

$$
\bar{\beta}\left(r,0\right):=\max\{r,k_\beta\bar{\beta}\left(r,1\right)\}\qquad k_\beta\in\mathbb{R}_{>1}.
\tag{49}
$$

The second term in the right hand side of equation (45), can be bounded by the following inequality $\forall j\in\mathbb{Z}_{[k-N,k-1]}$

$$
\begin{aligned}
\gamma_1\left(\|\boldsymbol{w}_j-\hat{\boldsymbol{w}}_{j|k}\|\right) &\leq \gamma_1\left(\|\boldsymbol{w}\|+\|\hat{\boldsymbol{w}}_{j|k}\|\right)\\
&\leq \gamma_1\left(\|\boldsymbol{w}\|+\underline{\gamma}_w^{-1}\left(\frac{3}{N}\bar{\gamma}_p\left(|x_{k-N}-\bar{x}_{k-N}|\right)\right)+\underline{\gamma}_w^{-1}\left(3\,\bar{\gamma}_w\left(\|\boldsymbol{w}\|\right)\right)+\underline{\gamma}_w^{-1}\left(3\,\bar{\gamma}_v\left(\|\boldsymbol{v}\|\right)\right)\right).
\end{aligned}
\tag{50}
$$

Recalling Assumption 2.2 and properties (1), the reader can verify the following inequality

$$
\gamma_1\left(\|\boldsymbol{w}_j-\hat{\boldsymbol{w}}_{j|k}\|\right)\leq\frac{c_1 3^{\alpha_1}|P_\infty^{-1}|^{\alpha_1}}{N^{\alpha_1}}\,|x_{k-N}-\bar{x}_{k-N}|^{a\alpha_1}+c_1 3^{\alpha_1}\bar{\gamma}_v^{\alpha_1}\left(\|\boldsymbol{v}\|\right)+\gamma_1\left(6\|\boldsymbol{w}\|\right)+\gamma_1\left(6\underline{\gamma}_w^{-1}\left(3\bar{\gamma}_w\left(\|\boldsymbol{w}\|\right)\right)\right).
\tag{51}
$$

In an equivalent manner, a bound for the third term in the right hand side of equation (45) can be found

$$
\gamma_2\left(\|\boldsymbol{v}_j-\hat{\boldsymbol{v}}_{j|k}\|\right)\leq\frac{c_2\,3^{\alpha_2}\,|P_\infty^{-1}|^{\alpha_2}}{N^{\alpha_2}}|x_{k-N}-\bar{x}_{k-N}|^{a\alpha_2}+c_2\,3^{\alpha_2}\bar{\gamma}_w^{\alpha_2}\left(\|\boldsymbol{w}\|\right)+\gamma_2\left(6\|\boldsymbol{v}\|\right)+\gamma_2\left(6\underline{\gamma}_v^{-1}\left(3\bar{\gamma}_v\left(\|\boldsymbol{v}\|\right)\right)\right).
\tag{52}
$$

Once the upper bounds for the three terms of (45) were found, defining $\zeta:=\max\{p,a\alpha_1,a\alpha_2\}$, $\eta:=\min\{q,\alpha_1,\alpha_2\}$ and $\rho:=\max\{p,\alpha_1,\alpha_2\}$, equation (45) can be rewritten as follows

$$
\begin{aligned}
|x_k-\hat{x}_{k|k}| &\leq \frac{|x_{k-N}-\bar{x}_{k-N}|^\zeta}{N^\eta}\,\mathbb{C}_P^\rho\left(c_\beta\,18^p+\left(c_1\,3^{\alpha_1}+c_2\,3^{\alpha_2}\right)\,\lambda_{\min}^{\alpha_1}\left(P_0^{-1}\right)+c_\beta\,2^p\right)+\\
&\quad\frac{c_\beta\,18^p}{|P_0^{-1}|}\,\bar{\gamma}_w^{\frac{p}{a}}\left(\|\boldsymbol{w}\|\right)+c_2\,3^{\alpha_2}\bar{\gamma}_w^{\alpha_2}\left(\|\boldsymbol{w}\|\right)+\gamma_1\left(6\|\boldsymbol{w}\|\right)+\gamma_1\left(6\underline{\gamma}_w^{-1}\left(3\bar{\gamma}_w\left(\|\boldsymbol{w}\|\right)\right)\right)+\\
&\quad\frac{c_\beta\,18^p}{|P_0^{-1}|}\,\bar{\gamma}_v^{\frac{p}{a}}\left(\|\boldsymbol{v}\|\right)+c_1\,3^{\alpha_1}\bar{\gamma}_v^{\alpha_1}\left(\|\boldsymbol{v}\|\right)+\gamma_2\left(6\|\boldsymbol{v}\|\right)+\gamma_2\left(6\underline{\gamma}_v^{-1}\left(3\bar{\gamma}_v\left(\|\boldsymbol{v}\|\right)\right)\right).
\end{aligned}
\tag{53}
$$

sinc(i) Research Institute for Signals, Systems and Computational Intelligence (sinc.unl.edu.ar)
Nestor N. Deniz, M. Murillo, G. Sanchez & L. Giovanini; "Robust stability of moving horizon estimation for non-linear systems with bounded disturbances using adaptive arrival cost"
IET Control Theory \& Applications, Vol. 14, No. 18, pp. 2879-2888, 2020.

This equation can be written as follows

$$|x_k - \hat{x}_{k|k}| \leq \bar{\beta}\left(|x_{k-N} - \bar{x}_{k-N}|, N\right) + \phi_w\left(\|\boldsymbol{w}\|\right) + \phi_v\left(\|\boldsymbol{v}\|\right) \qquad \forall k \in \mathbb{Z}_{[0,N-1]}. \tag{54}$$

by defining the following functions as follows

$$\bar{\beta}\left(|x_{k-N} - \bar{x}_{k-N}|,\ N\right) := \frac{|x_{k-N} - \bar{x}_{k-N}|^{\zeta}}{N^{\eta}}\ \left(\mathbb{C}_P^{\rho}\left(c_\beta\ 18^p + \lambda_{\min}^{\alpha_1}\left(P_0^{-1}\right)\left(c_1\ 3^{\alpha_1} + c_2\ 3^{\alpha_2}\ \right)\right) + c_\beta\ 2^p\right), \tag{55}$$

$$\phi_w\left(\|\boldsymbol{w}\|\right) := \frac{c_\beta\ 18^p}{|P_0^{-1}|}\ \bar{\gamma}_w^{\frac{p}{a}}\left(\|\boldsymbol{w}\|\right) + c_2\ 3^{\alpha_2}\bar{\gamma}_w^{\alpha_2}\left(\|\boldsymbol{w}\|\right) + \gamma_1\left(6\|\boldsymbol{w}\|\right) + \gamma_1\left(6\underline{\gamma}_w^{-1}\left(3\bar{\gamma}_w\left(\|\boldsymbol{w}\|\right)\right)\right), \tag{56}$$

$$\phi_v\left(\|\boldsymbol{v}\|\right) := \frac{c_\beta\ 18^p}{|P_0^{-1}|}\ \bar{\gamma}_v^{\frac{p}{a}}\left(\|\boldsymbol{v}\|\right) + c_1\ 3^{\alpha_1}\bar{\gamma}_v^{\alpha_1}\left(\|\boldsymbol{v}\|\right) + \gamma_2\left(6\|\boldsymbol{v}\|\right) + \gamma_2\left(6\underline{\gamma}_v^{-1}\left(3\bar{\gamma}_v\left(\|\boldsymbol{v}\|\right)\right)\right). \tag{57}$$

Functions ϕ_w and ϕ_v only depend on the characteristic of the noises w and v, stage costs $\ell\left(w,v\right)$ (through its bounds $\underline{\gamma}_w, \underline{\gamma}_v, \bar{\gamma}_w$ and $\bar{\gamma}_v$), the functions γ_1 and γ_2 (used to measure *i*-IOSS property), and the initial value of the prior weight matrix P_0, therefore they only affect the steady state value of the residual defining a bounded set $\mathcal{X}(w,v) \in \mathcal{X}$ where the estimated states $\hat{x}_k$ will remain.

On the other hand, the function $\bar{\beta}$ depends on the error at the beginning of the estimation window $(x_{k-N} - \bar{x}_{k-N})$, stage costs $\ell\left(w,v\right)$ (through its bounds $\bar{\gamma}_w$ and $\bar{\gamma}_v$), the initial (P_0) and final (P_∞) values of the prior weight matrix through their eigenvalues and the size of the estimation horizon N.

The only term that can affect the stability of the estimates is β, since it depends on the error at the beginning of the estimation window $(x_{k-N} - \bar{x}_{k-N})$, functions γ_1 and γ_2 (used to define IOSS) and the behaviour of the updating mechanism of the arrival cost parameters (P_k). Then, we want to determinate the minimum horizon length $\mathbb{N}_{min}$ required to accomplish a decreasing rate $\delta \in (0,1)$ such that $\forall r, s:$ $\bar{\beta}(r,s) \geq \beta(r,s)$, which is given by

$$N \geq \mathbb{N}_{min} = \left(\delta^{\zeta}\ e_{\max}^{\zeta-1}\ \mathbb{C}_P^{\rho}\left(c_\beta 18^p + \lambda_{\min}^{\alpha_1}\left(P_0^{-1}\right)\left(c_1 3^{\alpha_1} + c_2 3^{\alpha_2}\right) + c_\beta 2^p\right)\right)^{\frac{1}{\eta}}. \tag{58}$$

Adopting an estimator with a window length $N \geq \mathbb{N}_{min}$, $\bar{\beta}(r,s)$ is bounded by

$$\bar{\beta}\left(\delta r,\ N\right) \leq \left(\frac{\mathbb{N}_{min}}{N}\right)^{\eta} r, \tag{59}$$

and the effects of the initial conditions $|x_0 - \bar{x}_0| > 0$ will vanish with a decreasing rate δ. As $k \to \infty$, the estimates $\hat{x}_{k|k}$ will entry to the bounded set $\mathcal{X}_I\left(w,v\right) \in \mathcal{X}$ defined by the noises of the system

$$\mathcal{X}_I\left(w,v\right) := \left\{|x_k - \hat{x}_{k|k}| \leq 2\left(1+\mu\right)\left(\phi_w\left(\|\boldsymbol{w}\|\right) + \phi_v\left(\|\boldsymbol{v}\|\right)\right)\right\} \qquad \mu \in \mathbb{R}_{>0}. \tag{60}$$

This fact completes the proof of Theorem 1. □