# Adaptive polytopic estimation for nonlinear systems under bounded disturbances using moving horizon ⋆

Nestor Deniz [a], Marina Murillo [a], Guido Sanchez [a], Leonardo Giovanini [a]

[a] *Instituto de Investigacion en Senales, Sistemas e Inteligencia Computacional, sinc(i), UNL, CONICET, Ciudad Universitaria UNL, 4to piso FICH, (S3000) Santa Fe, Argentina*



**Abstract**

This paper introduces an adaptive polytopic estimator design for nonlinear systems under bounded disturbances combining moving horizon and dual estimation techniques. It extends the moving horizon estimation results for *LTI* systems to polytopic *LPV* systems. The design and necessary conditions to guarantee the robust stability and convergence to the true state and parameters for the case of bounded disturbances and convergence to the true system and state are given for the vanishing disturbances.

*Key words:* Adaptive polytopic observer, Moving horizon estimation, quasi-LPV systems, Nonlinear systems.

## 1 Introduction

Accurate information of states, parameters and disturbances is essential for effective real-time operation of any system. Many of its relevant variables are often not measurable or too expensive to measure on line. A cost effective approach is to employ estimation techniques to obtain the required information from measurements of other variables and a mathematical model of the

⋆ The material in this paper was not presented at any conference.

*Email addresses:* ndeniz@sinc.unl.edu.ar (Nestor Deniz), mmurillo@sinc.unl.edu.ar (Marina Murillo), gsanchez@sinc.unl.edu.ar (Guido Sanchez), lgiovanini@sinc.unl.edu.ar (Leonardo Giovanini).

system. Linear and nonlinear estimation have been an active researcher field during the past several decades (Patwardhan et al. 2012). Linear estimation methods use a simpler representation of the system and can provide acceptable performance only around an operating point and the steady state operational conditions. However, as nonlinearities in the system dynamics become dominants, the performance of linear approaches deteriorates and the estimation algorithms will not necessarily converge to an accurate solutions. Although optimal state estimation solutions for linear systems exists, nonlinear estimation algorithms suffer from generating near-optimal solutions. Consequently, research of nonlinear estimation and filtering problems remains a challenging research area.

Countless studies have been conducted in the literature to address and analyse nonlinear estimation problems. These methods can be broadly categorized into (Patwardhan et al. 2012): i) linearization methods (Kushner 1977); ii) approximation methods (Beneš 1981); iii) Bayesian recursive methods (Doucet et al. 2001); iv) moment methods (Crisan et al. 1998); and v) higher dimensional nonlinear filter methods (Arasaratnam & Haykin 2009). However, there exist some approaches that approximate the nonlinear behaviour of systems with a linear parameter–varying (LPV) models (Shamma & Athans 1991, Shamma & Cloutier 1993).

LPV systems are linear systems with matrices depending on time-varying parameters that can evolve over wide operating ranges (Apkarian et al. 1995). These parameters, called *scheduling variables*, depend on exogenous signals that can be measured. When the bounds of these signals are known, the LPV model can be reformulated into a convex linear combination of linear time-invariant (Leith & Leithead 2000). If the scheduling variables are functions of endogenous signals such as states, inputs or outputs of the system instead of exogenous signals, $LPV$ system describes a large class of nonlinear systems (Tóth et al. 2011). The most common technique to obtain an LPV system is the polytopic approach, where the system depends affinely on a time-varying parameter vector that evolves within a polytopic set. In practical situations they could be inaccessible by the fact that scheduling variables are functions of the system states (Theilliol & Aberkane 2011).

In the polytopic LPV observer design, trust full knowledge of the scheduling variables is of paramount importance, because this information is needed to design the observer. Many researchers have proposed solutions to this problem in the polytopic framework. LPV observers with unmeasured scheduling parameters can be designed using proportional observer (Ichalal et al. 2016), proportional-integral observer (Aouaouda et al. 2013), generalized dynamic observer (Gao et al. 2016, Osorio-Gordillo et al. 2016) and adaptive observer (Bezzaoucha et al. 2013, 2018) framework, respectively.

The main contribution of this paper is the design and analysis of a robust estimator for nonlinear systems under bounded disturbances combining quasi-$LPV$ models and dual estimation using a receding horizon framework. The proposed algorithm simultaneously estimates the mixing parameters and the states using a dual estimation approach within a multiple iteration scheme that improve the performance of the estimation at each sample. The conditions to guarantee the robust stability and a convergence to the true system and states for the case of vanishing disturbances are derived. To achieve these results is crucial that the prior weighting in the cost function and the length of the estimation horizon are properly chosen. The assumption on the prior weighting can be verified a prior design. 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 results and shows its connections with previous results. The stability and convergence to the true state in the dual iteration is discussed in the initial part of section. Then, the robust regional stability and convergence to the true state and parameters of the estimator are analysed. In section 4 two simple examples are discussed to illustrate the concepts and to show the difference with the estate of the art. 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 $[a, b] \subseteq \mathbb{R}$, and $\mathbb{Z}_{\geq a}$ denotes the set of integers greater or equal to $a$. Boldface symbols denote sequences of finite or infinite length, i.e., $\boldsymbol{w} := \{w_{k_1}, \ldots, w_{k_2}\}$ for some $k_1, k_2 \in \mathbb{Z}_{\geq 0}$ and $k_1 < k_2$, respectively. We denote $x_{j|k}$ as the element of the sequence $\boldsymbol{x}$ given at time $k \in \mathbb{Z}_{\geq 0}$ and $j \in [k_1, k_2]$. By $|x|$ we denote the Euclidean norm of a vector $x \in \mathbb{R}^n$. Let $\|\boldsymbol{x}\| := \sup_{k \in \mathbb{Z}_{\geq 0}} |x_k|$ denote the supreme norm of the sequence $\boldsymbol{x}$ and $\|\boldsymbol{x}\|_{[a,b]} := \sup_{k \in \mathbb{Z}_{[a,b]}} |x_k|$ . A function $\gamma : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class $\mathscr{K}$ if $\gamma$ is continuous, strictly increasing and $\gamma(0) = 0$ . If $\gamma$ is also unbounded, it is of class $\mathscr{K}_\infty$. A function $\zeta : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class $\mathscr{L}$ 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 $\mathscr{KL}$ if $\beta(\cdot, k)$ is of class $\mathscr{K}$ for each fixed $k \in \mathbb{Z}_{\geq 0}$, and $\beta(r, \cdot)$ of class $\mathscr{L}$ for each fixed $r \in \mathbb{R}_{\geq 0}$.

The following inequalities hold for all $\beta \in \mathscr{KL}$, $\gamma \in \mathscr{K}$ and $a_j \in \mathbb{R}_{\geq 0}$ with $j \in$

$\mathbb{Z}_{[1,n]}$

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

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

**Bounded sequences:** A sequence $\boldsymbol{w}$ is bounded if $\|\boldsymbol{w}\|$ is finite. The set of bounded sequences $\boldsymbol{w}$ is denoted as $\mathscr{W}\left(w_{\max}\right) := \{w : \boldsymbol{w} \leq w_{\max}\}$ for some $w_{\max} \in \mathbb{R}_{\geq 0}$.

**Convergent sequences:** 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

$$\mathscr{C}_w := \{\boldsymbol{w} \in \mathscr{W}\left(w_{\max}\right) | \boldsymbol{w} \text{ is convergent}\}.$$

Analogously, the sequence $\boldsymbol{v}$ and $\mathscr{C}_v$ can be defined in similar way.

### *2.2 Problem statement*

Let us consider a nonlinear discrete-time system with the following behaviour

$$\begin{aligned} x_{k+1} =& f\left(x_k, w_k,\right) \qquad x_0 = \mathtt{x}_0,\ \forall k \in \mathbb{Z}_{\geq 0}, \\ y_k =& h\left(x_k\right) + v_k \end{aligned} \tag{2}$$

where $x_k \in \mathscr{X} \subset \mathbb{R}^{n_x}$ is the system state, $w_k \in \mathscr{W} \subset \mathbb{R}^{n_x}$ is the additive process disturbance, $y_k \in \mathscr{Y} \subset \mathbb{R}^{n_p}$ is the system measurements and $v_k \in \mathscr{V} \subset \mathbb{R}^{n_p}$ is the measurement noise. The sets $\mathscr{X}, \mathscr{W}, \mathscr{Y}$ and $\mathscr{V}$ are known compact and convex with the null vector $\mathbf{0}$ in their interior. In the following we assume that $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_m} \to \mathbb{R}^{n_x}$ is at least $C^1$ and locally Lipschitz on $x_k$ and $h : \mathbb{R}^{n_x} \to \mathbb{R}^{n_p}$ is continuous. Finally, the solution of system (2) at time $k$ is denoted by $x(k, x_0, \boldsymbol{w}, \boldsymbol{d})$, with initial condition $\mathtt{x}_0$ and disturbance sequence $\boldsymbol{w}$. Furthermore, the initial conditions $x_0$ and $\alpha_0$ are unknown, but priors knowledge $\bar{x}_0$ and $\bar{\alpha}_0$ are assumed to be available and their errors are assumed to be bounded, i.e., $\bar{x}_0 \in \mathscr{X}_0 := \{\bar{x}_0 : |x_0 - \bar{x}_0| \leq e_{x\,max}\}$ such that $\mathscr{X}_0 \subseteq \mathscr{X}$ and $\bar{\alpha}_0 \in \mathscr{A}_0 := \{\bar{\alpha}_0 : |\alpha_0 - \bar{\alpha}_0| \leq e_{\alpha\,max}\}$ such that $\mathscr{A}_0 \subseteq \mathscr{A}$, respectively.

The solution of the estimation problem aims to find at time $k$ an estimate $\hat{x}_{k|k}$ of the current state $x_k$ using a moving horizon estimator (*MHE*). At each sampling time $k$ the only information available are the previous $N$ measurements $\boldsymbol{y} := \{y_{k-N}, \ldots, y_k\}$ and a matrix $G(x_k, w_k) \in \Omega(\mathscr{A})$, where $\mathscr{A}$ denotes a polytopic set of matrices such that

$$A_k = \sum_{i=1}^{q} \alpha_{i,k} A_i, \quad \sum_{i=1}^{q} \alpha_{i,k} C_i \tag{3}$$

with $\mathscr{A}$ the unit simplex

$$\mathscr{A} := \left\{ \sum_{i=1}^{q} \alpha_{i,k} = 1, \alpha_{i,k} \geq 0 \right\} \tag{4}$$

Then, any property ensured for the uncertain $LPV$ model

$$\begin{aligned} x_{k+1} &= \sum_{i=1}^{q} \alpha_{i,k} A_i x_k + w_k + d_k, \\ y_k &= \sum_{i=1}^{q} \alpha_{i,k} C_i x_k + v_k, \end{aligned} \tag{5}$$

holds true also for the nonlinear system (2) (Angelis 2003). Therefore, in this work we propose a moving horizon estimation algorithm to simultaneously estimate the state of the system $\hat{x}_{k|k}$ and the mixing parameter of the $LPV$ model $\hat{\alpha}_{k|k}$. The optimization problem to be solved at each sampling time is the following

$$\min_{\hat{\boldsymbol{x}}_{k-N|k}, \hat{\boldsymbol{w}}, \hat{\boldsymbol{d}}, \hat{\boldsymbol{\alpha}}_k, \hat{\boldsymbol{w}}_{\boldsymbol{\alpha}}} \begin{aligned} \Psi_{x,\alpha} := \Gamma_{k-N}\left(\hat{x}_{k-N|k}\right) + \sum_{j=k-N}^{k} \ell\left(\hat{w}_{j|k}, \hat{v}_{j|k}, \hat{w}_{\alpha_{j|k}}, \hat{d}_{j|k}\right) \\ + \Lambda_k(\hat{\alpha}_{k-N|k}) \end{aligned}$$

$$\text{s.t.} \begin{cases} \hat{\alpha}_{j+1|k} = \hat{\alpha}_{j|k} + \hat{w}_{\alpha_{j|k}} & j \in \mathbb{Z}_{[k-N,k-1]}, \\ \hat{x}_{j+1|k} = \sum_{i=1}^{q} \hat{\alpha}_{i,k|k} A_i \hat{x}_{j|k} + \hat{w}_{j|k} + \hat{d}_{j|k}, & \\ y_j = \sum_{i=1}^{q} \hat{\alpha}_{i,k|k} C_i \hat{x}_{j|k} + \hat{v}_{j|k} & j \in \mathbb{Z}_{[k-N,k]}, \\ \sum_{i=1}^{q} \hat{\alpha}_{i,k|k} = 1, & \\ \hat{\alpha}_{i,k|k} \geq 0 & i \in \mathbb{Z}_{[1,q]}, \\ \hat{x}_{j|k} \in \mathscr{X}, \ \hat{w}_{j|k} \in \mathscr{W}, \hat{w}_{\alpha} \in \mathscr{W}_{\alpha}, \ \hat{v}_{j|k} \in \mathscr{V}, \ \hat{d}_{j|k} \in \mathscr{D}. & \end{cases} \tag{6}$$

where $\hat{x}_{j|k}$ is the optimal estimated, $\hat{w}_{j|k}$ is the optimal process noise estimate and $\hat{\alpha}_{j|k}$ is the optimal mixing parameter and $\hat{w}_{\alpha_{j|k}}$ the noise associated to it at sample $k-j \quad j = 0, 1, \ldots, N$ based on measurements $y_{k-j}$ available at time $k$. The process noise $\hat{\boldsymbol{w}} := \left\{\hat{w}_{k-N-1|k}, \ldots, \hat{w}_{k-1|k}\right\}$, the mixing parameters $\alpha := \left[\alpha_{1,k|k}, \ldots, \alpha_{q,k|k}\right]^T$, $\hat{\boldsymbol{w}}_{\boldsymbol{\alpha}} := \left\{\hat{w}_{\alpha_{1,k|k}}, \ldots, \hat{w}_{\alpha_{q,k|k}}\right\}$ and $\hat{x}_{k-N|k}$ are the optimization variables. The stage cost $\ell\left(\hat{w}_{j|k}, \hat{v}_{j|k}, \hat{w}_{\alpha_{j|k}}, \hat{d}_{j|k}\right)$ penalizes the estimated process noise sequence $\hat{\boldsymbol{w}}_{j|k}$ and the estimation residuals $\hat{\boldsymbol{v}}_{j|k} = y_j - h\left(\hat{x}_{j|k}\right)$, while $\Gamma_{k-N}\left(\hat{x}_{k-N|k}\right)$ and $\Lambda\left(\hat{\alpha}_{k-N|k}\right)$ are the prior weights that penalizes the prior estimates $\hat{x}_{k-N|k}$ and $\hat{\alpha}_{k-N|k}$.

The robust stability of estimator (6) can be achieved by combining a suitable choice of the stage cost $\ell\left(\hat{w}_{j|k}, \hat{v}_{j|k}, \hat{d}_{j|k}\right)$ and the time–varying prior weights

$$\begin{aligned} \Gamma_{k-N|k}\left(\hat{x}_{k-N|k}\right) &= \mid \hat{x}_{k-N|k} - \bar{x}_{k-N} \mid_{P^{-1}_{x,k-N|k}}, \\ \Lambda_{k-N|k}\left(\hat{w}_{\alpha,k-N|k}\right) &= \mid \hat{\alpha}_{k-N|k} - \bar{\alpha}_{k-N} \mid_{P^{-1}_{\alpha,k-N|k}}, \end{aligned} \tag{7}$$

whose parameters $\left(P^{-1}_{x,k-N|k}, \bar{x}_{k-N}, P^{-1}_{\alpha,k-N|k}, \bar{\alpha}_{k-N}\right)$ are recursively updated using the information available at time $k$. In this approach, the prior weight matrix $P_{*,k-N|k}$ are given by (Sánchez et al. 2017)

$$\begin{aligned} \epsilon_{k-N} &= y_{k-N} - \hat{y}_{k-N|k}, \\ N_{*,k} &= \left[1 + \hat{*}^{T}_{k-N|k-1} P_{*,k-N-1} \hat{*}_{*,k-N|k-1}\right] \frac{\sigma}{|\epsilon_{k-N}|_2^2} \\ \theta_{*,k} &= 1 - \frac{1}{N_{*,k}}, \\ W_{*,k} &= \left[I - \frac{P_{*,k-N-1} \hat{*}_{k-N|k-1} \hat{*}^{T}_{k-N|k-1}}{1 + \hat{*}^{T}_{k-N|k-1} P_{*,k-N-1} \hat{*}_{k-N|k-1}}\right] P_{*,k-N-1}, \\ P_{*,k-N} &= \begin{cases} \frac{1}{\theta_{*,k}} W_{*,k} & \text{if } \frac{1}{\theta_{*,k}} Tr\left(W_{*,k}\right) \leq c, \\ W_{*,k} & \text{otherwise}, \end{cases} \end{aligned} \tag{8}$$

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

$$\begin{aligned} \bar{x}_{k-N} &= \hat{x}_{k-N|k-1}, \\ \bar{\alpha}_{k-N} &= \hat{\alpha}_{k-N|k-1}. \end{aligned} \tag{9}$$

### *2.3 Dual estimation formulation*

The joint estimator simultaneously estimates states and mixing parameters. For systems with many parameters, augmenting the state vector can cause a significant increase to the state dimension. This may be problematic as the dimension of the state vector grows, the errors accumulate and the convexity of the optimization problem is lost. To overcome this problem, a dual estimation setup is introduced: the estimation problem (6) solves separately the state estimation problem (assuming that mixing parameters $\alpha$ remains constant) and the model identification problem (assuming that estimated states $\hat{x}_k$ remains constant) at each sampling time. The problems to be solved iteratively are

$$\min_{\hat{\boldsymbol{x}}_{k-N|k},\hat{\boldsymbol{w}}_{\boldsymbol{x}}} \Psi_x := \Gamma_{k-N}\left(\hat{x}_{k-N|k}\right) + \sum_{j=k-N}^{k} \ell\left(\hat{w}_{j|k}, \hat{v}_{j|k}\right)$$
$$\text{s.t.} \begin{cases} \hat{x}_{k-N|k} = \bar{x}_{k-N} + \hat{w}_{k-N|k} & \\ \hat{x}_{j+1|k} = \sum_{i=1}^{q} \alpha_{i,k} A_i \hat{x}_{j|k} + \hat{w}_{x\ j|k} & j \in \mathbb{Z}_{[k-N,k-1]}, \\ y_j = \sum_{i=1}^{q} \alpha_{i,k} C_i \hat{x}_{j|k} + \hat{v}_{j|k} & j \in \mathbb{Z}_{[k-N,k]}, \\ \hat{x}_{j|k} \in \mathscr{X},\ \hat{w}_{j|k} \in \mathscr{W},\ \hat{v}_{j|k} \in \mathscr{V}. & \end{cases} \tag{10}$$

where the decision variables are $\hat{x}_{k-N|k}$ and $\hat{w}_{k-j|k} \quad j := 1, 2, \ldots, N$ and

$$\min_{\boldsymbol{\alpha},\hat{\boldsymbol{w}}_{\boldsymbol{\alpha}}} \Psi_\alpha := \Lambda_{k-N}\left(\hat{\alpha}_{k-N|k}\right) + \sum_{j=k-N}^{k} \ell\left(\hat{d}_{j|k}, \hat{v}_{j|k}, \hat{w}_{\alpha_{j|k}}\right)$$
$$\text{s.t.} \begin{cases} \alpha_{k-N|k} = \bar{\alpha}_{k-N} + \hat{w}_{\alpha_{k-N|k}} & \\ \alpha_{j+1|k} = \alpha_{j|k} + \hat{w}_{\alpha_{j|k}} & j \in \mathbb{Z}_{[k-N,k-1]} \\ x_{j+1|k} = \sum_{i=1}^{q} \alpha_{i,j|k} A^i x_{j|k} + \hat{d}_{j|k}, & \\ y_j = \sum_{i=1}^{q} \alpha_{i,j|k} C^i x_{j|k} + \hat{v}_{j|k} & j \in \mathbb{Z}_{[k-N,k]}, \\ \sum_{i=1}^{q} \alpha_{i,j|k} = 1 & \\ \alpha_{i,j|k} \geq 0 & j \in \mathbb{Z}_{[1,q]} \\ \hat{d}_{j|k} \in \mathscr{D}, \hat{w}_\alpha \in \mathscr{W}_\alpha,\ \hat{v}_{j|k} \in \mathscr{V}. & \end{cases} \tag{11}$$

where the decision variables are $\hat{w}_{\alpha,k-j|k} \quad j := 0, 1, \ldots, N$ and $\hat{\alpha}_{k-N|k}$. Problems (10) and (11) are solved iteratively several times for the same sampling-time. The main novelty of the proposed algorithm is that an improvement in the state estimation and model identification can be guaranteed for a certain number of iterations when some assumptions are fulfilled. Moreover, the number of iterations can be computed offline.

The sequence $P_{k|k} \quad k \geq 0$ is positive definite, it is decreasing in norm and it is bounded. The proof of these properties follows similar steps as in Sánchez et al. (2017).

**Assumption 1** *The prior weighting* $\Gamma_{k-N}\left(\hat{x}_{k-N|k}\right)$ *is a continuous function* $\mathbb{R}^n \to \mathbb{R}$ *lower bounded by* $\underline{\gamma}_p(\cdot) \in \mathscr{K}_\infty$ *and upper bounded by* $\bar{\gamma}_p(\cdot) \in \mathscr{K}_\infty$ *such that*

$$\underline{\gamma}_p\left(\left|\hat{x}_{k-N|k} - \bar{x}_{k-N}\right|\right) \leq \Gamma_{k-N}\left(\left|\hat{x}_{k-N|k} - \bar{x}_{k-N}\right|\right) \leq \bar{\gamma}_p\left(\left|\hat{x}_{k-N|k} - \bar{x}_{k-N}\right|\right) \tag{12}$$

*for all* $\hat{x} \in \mathscr{X}$ *and*

$$\underline{\gamma}_p(r) \geq \underline{c}_p\, r^a, \quad \bar{\gamma}_p(r) \leq \bar{c}_p\, r^a \tag{13}$$

*where* $0 \leq \underline{c}_p \leq \bar{c}_p$ *and* $a \in R_{\geq 1}$. *Moreover, if the arrival cost is updated using equation* (8)*, the bounds* $\underline{\gamma}_p$ *and* $\bar{\gamma}_p$ *are bounded by*

$$\underline{\gamma}_p(r) \geq |P_0^{-1}|\, r^a, \quad \overline{\gamma}_p(r) \leq |P_\infty^{-1}|\, r^a \tag{14}$$

**Definition 1** *The system* (2) *is incrementally input/output-to-state stable if there exist some functions* $\beta \in \mathscr{KL}$ *and* $\gamma_1, \gamma_2 \in \mathscr{K}$ *such that for every two initial states* $z_1$, $z_2 \in \mathbb{R}^n$*, and any two disturbances sequences* $\boldsymbol{w_1}, \boldsymbol{w_2}$ *the following holds for all* $k \in \mathbb{Z}_{\geq 0}$*:*

$$\begin{aligned} |x(k, z_1, \boldsymbol{w_1}) - x(k, z_2, \boldsymbol{w_2})| \leq \max\Big\{&\beta(|z_1 - z_2|, k), \gamma_1\left(\|\boldsymbol{w_1} - \boldsymbol{w_2}\|_{[0,k-1]}\right), \\ &\gamma_2\left(\|\boldsymbol{y_1} - \boldsymbol{y_2}\|_{[0,k-1]}\right)\Big\} \\ \leq\ &\beta(|z_1 - z_2|, k) + \gamma_1\left(\|\boldsymbol{w_1} - \boldsymbol{w_2}\|_{[0,k-1]}\right) + \\ &\gamma_2\left(\|\boldsymbol{y_1} - \boldsymbol{y_2}\|_{[0,k-1]}\right) \end{aligned} \tag{15}$$

for all $k$ (Sontag & Wang 1997).

**Assumption 2** *The function* $\beta(r, s) \in \mathscr{KL}$ *and satisfies the following inequality:*

$$\beta(r, s) \leq c_\beta r^p s^{-q} \tag{16}$$

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

**Assumption 3** *The stage cost* $\ell(\cdot) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ *is a continuous function bounded by* $\underline{\gamma}_w, \underline{\gamma}_v, \bar{\gamma}_w, \bar{\gamma}_v \in \mathscr{K}_\infty$ *such that the following inequalities are satisfied* $\forall w \in \mathscr{W}$ and $v \in \mathscr{V}$

$$\underline{\gamma}_w(\hat{w}) + \underline{\gamma}_v(\hat{v}) \leq \ell(\hat{w}, \hat{v}) \leq \bar{\gamma}_w(\hat{w}) + \bar{\gamma}_v(\hat{v}) \tag{17}$$

Functions $\gamma_1$ and $\gamma_2$ from Definition 1 are related with the bounds of stage cost $\bar{\gamma}_w, \underline{\gamma}_w, \bar{\gamma}_v$ and $\underline{\gamma}_v$ through the following inequalities

$$\gamma_1\left(6\underline{\gamma}_w^{-1}(r)\right) \leq c_1 r^{b_1},\ \gamma_2\left(6\underline{\gamma}_v^{-1}(r)\right) \leq c_2 r^{b_2} \tag{18}$$

for $c_1, c_2, b_1, b_2 > 0$.

**Assumption 4** *The prior weighting* $\Lambda_{k-N}\left(\hat{\alpha}_{k-N|k}\right)$ *is a continuous function* $\mathbb{R}^q \to \mathbb{R}$ *lower bounded by* $\underline{\gamma}_\Lambda(\cdot) \in \mathscr{K}_\infty$ *and upper bounded by* $\overline{\gamma}_\Lambda(\cdot) \in \mathscr{K}_\infty$ *such that*

$$\underline{\gamma}_\Lambda(|\hat{\alpha}_{k-N} - \bar{\alpha}_{k-N}|) \leq \Lambda_{k-N}(|\hat{\alpha}_{k-N} - \bar{\alpha}_{k-N}|) \leq \overline{\gamma}_\Lambda(|\hat{\alpha}_{k-N} - \bar{\alpha}_{k-N}|) \tag{19}$$

*where*

$$\begin{aligned} \underline{\gamma}_\Lambda\left(|\alpha_{k-N}-\bar{\alpha}_{k-N}|\right) &\geq \underline{c}_\Lambda|\alpha_{k-N}-\bar{\alpha}_{k-N}|^a \\ \overline{\gamma}_\Lambda\left(|\alpha_{k-N}-\bar{\alpha}_{k-N}|\right) &\leq \overline{c}_\Lambda|\alpha_{k-N}-\bar{\alpha}_{k-N}|^a \end{aligned} \tag{20}$$

*for some* $\underline{c}_\Lambda \in \mathbb{R}_{\geq 0}$, $\overline{c}_\Lambda \in \mathbb{R}_{\geq 0}$, $a \in \mathbb{R}_{>0}$, $\overline{c}_\Lambda > \underline{c}_\Lambda$.

In this work, we claim that the proposed estimator holds the property of being robust asymptotic stable, which is defined as follows.

**Definition 2** *Consider the system described by Equation* (11) *subject to disturbances* $\boldsymbol{w} \in \mathscr{W}(w_{\max})$ *and* $\boldsymbol{v} \in \mathscr{V}(v_{\max})$ *for* $w_{\max} \in \mathbb{R}_{\geq 0}$, $v_{max} \in \mathbb{R}_{\geq 0}$ *with prior estimate* $\bar{x}_0 \in \mathscr{X}(e_{\max})$ *for* $e_{\max} \in \mathbb{R}_{\geq 0}$. *The moving horizon state estimator given by Equation* (11) *is robustly asymptotically stable (RAS) if there exists functions* $\Phi \in \mathscr{KL}$ *and* $\pi_w$, $\pi_v \in \mathscr{K}$ *such that for all* $x_0 \in \mathscr{X}$, *all* $\bar{x}_0 \in \mathscr{X}_0(e_{\max})$, *the following is satisfied for all* $k \in \mathbb{Z}_{\geq 0}$

$$|x_k - \hat{x}_k| \leq \Phi\left(|x_0 - \bar{x}_0|, k\right) + \pi_w\left(\|\boldsymbol{w}\|_{[0,k-1]}\right) + \pi_v\left(\|\boldsymbol{v}\|_{[0,k-1]}\right) \tag{21}$$

We want to show that if system (10) is i-IOSS and Assumptions (1), (2) and (3) are fulfilled, then the proposed estimator with adaptive arrival cost is RGAS. Furthermore, if the process disturbance and measurement noise sequences are convergent (i.e., $\boldsymbol{w}, \boldsymbol{v} \in \mathscr{C}$), the estimation converges to the true state.

## 3 Theoretical properties

Now we are ready to derive the main results: $i$) the stability of the iterated dual estimation and $ii$ the robust asymptotic stability ($RAS$) of the proposed estimator with an estimation horizon $\mathscr{N}$ for nonlinear detectable systems under bounded disturbances.

### *3.1 Stability of the dual estimation iterations*

As stated formerly, problems (10)–(11) are solved sequentially within a dual estimation iteration for each sampling time. In the following, we will state the conditions required to achieve effectively a decreasing behaviour of the costs inside the dual estimation iteration. We have now all the necessary ingredients to enunciate the first theorem,

**Theorem 1** *The sequences of costs* $\{\Psi_x^1, \Psi_x^2, \ldots, \Psi_x^l\}$ *and* $\{\Psi_\alpha^1, \Psi_\alpha^2, \ldots, \Psi_\alpha^l\}$ *generated by the dual estimation iteration are decreasing if the number of it-*

*erations $l$ satisfies*

$$l \geq \log_2 \left( \frac{\epsilon \Psi_x^1 - \Gamma_{k-N} \left( \hat{x}_{k-N|k}^l \right)}{\Gamma_{k-N} \left( \hat{x}_{k-N|k}^m \right)} + 1 \right) + 1, \tag{22}$$

*where $\epsilon \in \mathbb{R}_{\geq 0}$ and*

$$\Gamma_{k-N} \left( \hat{x}_{k-N|k}^m \right) := \min_{i \in \mathbb{Z}_{[1,l]}} \left\{ \Gamma_{k-N} \left( \hat{x}_{k-N|k}^i \right) \right\}.$$

**Proof.** Let us consider the sequence of costs $\Psi_x^i(\cdot)$ and $\Psi_\alpha^i(\cdot)$ $\forall i \in \mathbb{Z}_{>1}$ generated in the $i$–th iteration of the optimization problems (10) and (11). Due to the optimality of the solutions the following inequalities are satisfied

$$\begin{aligned} \Psi_x^1(\cdot) &\geq \Psi_x^2(\cdot) \geq, \ldots, \geq \Psi_x^l(\cdot), \\ \Psi_\alpha^1(\cdot) &\geq \Psi_\alpha^2(\cdot) \geq, \ldots, \geq \Psi_\alpha^l(\cdot). \end{aligned} \tag{23}$$

Since any iteration takes into account both $\Psi_x^i(\cdot)$ and $\Psi_\alpha^i(\cdot)$, and due to the sequences are non increasing, we only need to prove the decreasing behaviour of only one of these sequences, let's say $\Psi_x^i(\cdot)$.

Defining the normalized cost

$$g(k,i) := \frac{\Psi_x^i(\cdot)}{\Psi_x^1(\cdot)} \quad \forall k, i \geq 1, \tag{24}$$

the necessary and sufficient conditions to guarantee its decrement along the dual iteration can be obtained using the Gronwall inequality (see Ames & Pachpatte (1997), Holte (2009)). It states that given any three non-negative sequences $y_n$, $f_n$ and $g_n$ that satisfy

$$y_n \leq f_n + \sum_{k=0}^{n} g_k \, y_k \qquad \forall n > 0, \tag{25}$$

they also verify

$$y_n \leq f_n + \sum_{k=0}^{n-1} f_k g_k \prod_{j=k+1}^{n-1} (1 + g_j) \qquad \forall n > 0. \tag{26}$$

Taking the sequences of costs $y_i = \Psi_x^i(\cdot)$, of arrival-costs $f_i = \Gamma_{k-N}(\cdot)$ and the normalized costs $g_i = g(k,i)$ for $n = l$ , which verify (25), Gronwall inequality (26) can be written as follows

$$\Psi_x^l \leq \Gamma_{k-N} \left( \hat{x}_{k-N|k}^l \right) + \sum_{i=1}^{l-1} \Gamma_{k-N} \left( \hat{x}_{k-N|k}^i \right) g(k,i) \prod_{j=i+1}^{l-1} (1 + g(k,j)). \tag{27}$$

Dividing by $\Psi_x^1$ we obtain

$$g\left(k,l\right) \leq \frac{\Gamma_{k-N}\left(\hat{x}^l_{k-N|k}\right) + \sum_{i=1}^{l-1} \Gamma_{k-N}\left(\hat{x}^i_{k-N|k}\right) g(k,i) \prod_{j=i+1}^{l-1} \left(1+g(k,j)\right)}{\Psi_x^1} = \epsilon,$$

which leads to

$$\sum_{i=1}^{l-1} \Gamma_{k-N}\left(\hat{x}^i_{k-N|k}\right) g(k,i) \prod_{j=i+1}^{l-1} \left(1+g(k,j)\right) = \epsilon \Psi_x^1 - \Gamma_{k-N}\left(\hat{x}^l_{k-N|k}\right). \quad (28)$$

Defining

$$\Gamma_{k-N}\left(\hat{x}^m_{k-N|k}\right) := \min_{i\in\mathbb{Z}_{[1,l]}} \left\{\Gamma_{k-N}\left(\hat{x}^i_{k-N|k}\right)\right\}$$

and recalling that $g\left(k,i\right)$ is a non-increasing sequence, $g\left(k,l\right) \leq g\left(k,i\right) \ \forall l \geq i$, equation (28) can be rewritten as follows

$$\Gamma_{k-N}\left(\hat{x}^m_{k-N|k}\right) g\left(k,l\right) \sum_{i=1}^{l-1} \prod_{j=i+1}^{l-1} \left(1+g\left(k,l\right)\right) < \epsilon \Psi_x^1 - \Gamma_{k-N}\left(\hat{x}^l_{k-N|k}\right). \quad (29)$$

Since

$$\sum_{i=1}^{l-1} \prod_{j=i+1}^{l-1} \left(1+g\left(k,l\right)\right) = \sum_{i=0}^{l-2} \left(1+g\left(k,l\right)\right)^i = \frac{\left(1+g\left(k,l\right)\right)^{l-1}-1}{g\left(k,l\right)}, \quad (30)$$

inequality (29) can be rewritten as follows

$$\Gamma_{k-N}\left(\hat{x}^m_{k-N|k}\right) g\left(k,l\right) \left(\frac{\left(1+g(k,l)\right)^{l-1}-1}{g(k,l)}\right) \overset{(30)}{<} \epsilon \Psi_x^1 - \Gamma_{k-N}\left(\hat{x}^l_{k-N|k}\right) \quad (31)$$

and finally $g\left(k,l\right)$ is bounded by

$$g\left(k,l\right) < \left(\frac{\epsilon \Psi_x^1 - \Gamma_{k-N}\left(\hat{x}^l_{k-N|k}\right)}{\Gamma_{k-N}\left(\hat{x}^m_{k-N|k}\right)} + 1\right)^{\frac{1}{l-1}} - 1 < 1. \quad (32)$$

Selecting a $l$ large enough, we can guarantee the decrement of sequence $g\left(k,i\right)$ and cost function $\Psi_x^i$ within the dual estimation iteration. Solving for inequality (32), an upper bound for the number of iterations is given by

$$l \geq \left\lceil \log_2 \left(\frac{\epsilon \Psi_x^1 - \Gamma_{k-N}\left(\hat{x}^l_{k-N|k}\right)}{\Gamma_{k-N}\left(\hat{x}^m_{k-N|k}\right)} + 1\right)\right\rceil + 1. \quad (33)$$

A conservative estimate of $l$ can be computed taking into account the worst case scenario

$$l \geq \left\lceil \log_2 \left(\mathscr{E} N \left(\bar{\gamma}_w\left(\|\boldsymbol{w}\|\right) + \bar{\gamma}_v\left(\|\boldsymbol{v}\|\right)\right) + 1\right)\right\rceil + 1, \quad (34)$$

where $\mathscr{E} := \epsilon / \Gamma_{k-N}\left(\hat{x}^{m}_{k-N|k}\right)$.

Inequalities (33) and (34) allow to compute the required value of $l$ to guarantee the costs decreasing within the dual estimation iteration. $\square$

**Remark 1** *Note that for the noiseless case, only one iteration is needed after the transient due to the uncertainty in the initial condition has vanished.*

### *3.2 Robust stability*

In the previous subsection it was shown that the sequence of cost decreases within the dual estimation iteration. At each sampling time, the model used by the estimator is replaced with the newly available until satisfied the stopping criteria. In the following paragraphs we will prove robust stability for the estimator under bounded disturbances and model uncertainty assuming that the system is $i$-IOSS. Moreover, if the length $N$ of the horizon of the estimator is larger than a certain value $\mathscr{N}$ that can be computed offline, the number of iterations $l$ is chosen according to equations (33) and (34), the effects of uncertainty in the initial condition vanish, as well as the disturbances due to model uncertainty. Besides, in the absence of process and measurement noises, states and model converges to the true ones.

**Theorem 2** *Consider an i-IOSS system* (2) *with disturbances* $\boldsymbol{w} \in \mathscr{W}(w_{\max})$, $\boldsymbol{v} \in \mathscr{V}(v_{\max})$. *Assume that the arrival cost weight matrix of the MHE problem* $\Gamma_{k-N}$ *is updated using the adaptive algorithm* (8). *Moreover, Assumptions 1, 2 and 3 are fulfilled and initial condition* $x_0$ *and* $\alpha_0$ *are unknown, but prior estimates* $\bar{x}_0 \in \mathscr{X}_0$ *and* $\bar{\alpha}_0 \in \mathscr{A}_0$ *are available. Then, the MHE estimator resulting from problems* (10)– (11) *is RAS.*

**Proof.** In order to proof stability for the estimator, we start comparing the costs of the first iteration and the resulting estimated state $\hat{x}_{k-N|k}$ for sampling time $k$

$$\begin{aligned}\Psi\left(\hat{x}_{k-N|k}, \hat{\alpha}, \left\{\hat{\boldsymbol{w}}_{j|k}\right\}, \left\{\hat{\boldsymbol{d}}_{j|k}\right\}\right) &= g\left(k,l\right)\Psi^1\left(\hat{x}^1_{k-N|k}, \hat{\alpha}^1, \left\{\hat{\boldsymbol{w}}^1_{j|k}\right\}, \left\{\hat{\boldsymbol{d}}^1_{j|k}\right\}\right)\\ &\le g\left(k,l\right)\Psi\left(x_{k-N}, \alpha, \left\{\boldsymbol{w}\right\}, \left\{\boldsymbol{0}\right\}\right)\end{aligned}$$

Note that if $\hat{\alpha} = \alpha$, the estimated model match with the system, therefore there is no model uncertainty, i.e. $\{\boldsymbol{d}\} = \{\boldsymbol{0}\}$. Replacing the sequence of estimated process noises $\{\hat{\boldsymbol{w}}\}$ by the true sequence $\{\boldsymbol{w}\}$, the only feasible solution is the true sequence of states $\{\boldsymbol{x}\}$, and due optimality the inequality is verified. By mean of Assumptions 1 - 4, the cost $\Psi\left(\cdot\right)$ is bounded by

$$
\begin{aligned}
&\underline{\gamma}_p\left(\left|\hat{x}_{k-N|k}-\bar{x}_{k-N}\right|\right)+\underline{\gamma}_\Lambda\left(\left|\hat{\alpha}_{k-N|k}-\bar{\alpha}_{k-N}\right|\right)+N\underline{\gamma}_w\left(\left|\hat{w}_{j|k}\right|\right)+\\
&\qquad\underline{\gamma}_\alpha\left(\left|\hat{w}_{\alpha,j|k}\right|\right)+N\underline{\gamma}_v\left(\left|\hat{v}_{j|k}\right|\right)+N\underline{\gamma}_d\left(\left|\hat{d}_{j|k}\right|\right)\leq\\
&\Big(\overline{\gamma}_p\left(\left|x_{k-N}-\bar{x}_{k-N}\right|\right)+\overline{\gamma}_\Lambda\left(\left|\alpha_{k-N}-\bar{\alpha}_{k-N}\right|\right)+N\overline{\gamma}_w\left(\|\boldsymbol{w}\|_{[k-N,k]}\right)+\\
&\qquad N\overline{\gamma}_\alpha\left(\|\boldsymbol{w}_\alpha\|_{[k-N,k]}\right)+N\overline{\gamma}_v\left(\|\boldsymbol{v}\|_{[k-N,k]}\right)\Big)\,g\left(k,l\right)\leq\\
&\overline{c}_p\left|x_{k-N}-\bar{x}_{k-N}\right|^a g\left(k,l\right)+\overline{c}_\Lambda\left|\alpha_{k-N}-\bar{\alpha}_{k-N}\right|g\left(k,l\right)+N\left(\overline{\gamma}_w\left(\|\boldsymbol{w}\|\right)+\right.\\
&\qquad\left.\overline{\gamma}_v\left(\|\boldsymbol{v}\|\right)+\overline{\gamma}_\alpha\left(\|\boldsymbol{w}_{\boldsymbol{\alpha}}\|\right)\right)g\left(k,l\right).
\end{aligned}
\tag{35}
$$

Solving for $|\hat{x}_{k-N|k}-\bar{x}_{k-N}|$ and using relations (1), we can write

$$
\begin{aligned}
\left|\hat{x}_{k-N|k}-\bar{x}_{k-N}\right|\leq\;&\underline{\gamma}_p^{-1}\left(\overline{c}_p\left|x_{k-N}-\bar{x}_{k-N}\right|^a g\left(k,l\right)\right)+\\
&\underline{\gamma}_p^{-1}\left(5\overline{c}_\Lambda\left|\alpha_{k-N}-\bar{\alpha}_{k-N}\right|g\left(k,l\right)\right)+\\
&\underline{\gamma}_p^{-1}\left(5N\overline{\gamma}_w\left(\|\boldsymbol{w}\|_{[k-N,k]}\right)g\left(k,l\right)\right)+\\
&\underline{\gamma}_p^{-1}\left(5N\overline{\gamma}_\alpha\left(\|\boldsymbol{w}_{\boldsymbol{\alpha}}\|_{[k-N,k]}\right)g\left(k,l\right)\right)+\\
&\underline{\gamma}_p^{-1}\left(5N\overline{\gamma}_v\left(\|\boldsymbol{v}\|_{[k-N,k]}\right)g\left(k,l\right)\right).
\end{aligned}
\tag{36}
$$

Using again Assumptions 1 - 4, one can write

$$
\begin{aligned}
\left|\hat{x}_{k-N|k}-\bar{x}_{k-N}\right|\leq\;&\left(\frac{5\overline{c}_p\left|x_{k-N}-\bar{x}_{k-N}\right|^a g\left(k,l\right)}{c_p}\right)^{1/a}+\\
&\left(\frac{5\overline{c}_\Lambda\left|\alpha_{k-N}-\bar{\alpha}_{k-N}\right|^a g\left(k,l\right)}{\underline{c}_p}\right)^{1/a}+\\
&\left(\frac{5N\overline{\gamma}_w\left(\|\boldsymbol{w}\|_{[k-N,k]}\right)g\left(k,l\right)}{\underline{c}_p}\right)^{1/a}+\\
&\left(\frac{5N\overline{\gamma}_\alpha\left(\|\boldsymbol{w}_{\boldsymbol{\alpha}}\|_{[k-N,k]}\right)g\left(k,l\right)}{\underline{c}_p}\right)^{1/a}+\\
&\left(\frac{5N\overline{\gamma}_v\left(\|\boldsymbol{v}\|_{[k-N,k]}\right)g\left(k,l\right)}{\underline{c}_p}\right)^{1/a}
\end{aligned}
\tag{37}
$$

From now on, we will drop the superindex $l$. Using Definition 1, the estimation error at time $k$, given the error at initial conditions ($k=0$), is bounded by

$$
\left|x_k-\hat{x}_{k|k}\right|\leq\beta\left(\left|x_0-\hat{x}_{0|k}\right|,k\right)+\gamma_1\left(\|\boldsymbol{w}-\hat{\boldsymbol{w}}\|_{[0,k-1]}\right)+\gamma_2\left(\|\boldsymbol{v}-\hat{\boldsymbol{v}}\|_{[0,k-1]}\right),
$$

and, assuming that $k = N$, we have

$$|x_k - \hat{x}_{k|k}| \leq \beta\left(|x_{k-N} - \hat{x}_{k-N|k}|, N\right) + \gamma_1\left(\|\boldsymbol{w} - \hat{\boldsymbol{w}}\|_{[k-N,k-1]}\right) + \gamma_2\left(\|\boldsymbol{v} - \hat{\boldsymbol{v}}\|_{[k-N,k-1]}\right). \tag{38}$$

To found a bound for the estimation error we need to bounds for the terms of the right hand of (38). Let us start with the first term using inequalities (1) such that the effect of estimation error at the beginning of the estimation window is bounded by

$$\begin{aligned}\beta\left(|x_{k-N} - \hat{x}_{k-N|k}|, N\right) &= \beta\left(|x_{k-N} - \bar{x}_{k-N} + \bar{x}_{k-N} - \hat{x}_{k-N|k}|, N\right)\\ &\leq \beta\left(|x_{k-N} - \bar{x}_{k-N}| + |\bar{x}_{k-N} - \hat{x}_{k-N|k}|, N\right)\\ &\leq \beta\left(|x_{k-N} - \bar{x}_{k-N}| + |\hat{x}_{k-N|k} - \bar{x}_{k-N}|, N\right)\\ &\leq \beta\left(2|x_{k-N} - \bar{x}_{k-N}|, N\right) + \beta\left(2|\hat{x}_{k-N|k} - \bar{x}_{k-N}|, N\right).\end{aligned} \tag{39}$$

Now, the first term of (39) can be rewritten using Assumption 2, and the second term with the use of (37)

$$\begin{aligned}\beta\left(|x_{k-N} - \hat{x}_{k-N|k}|, N\right) \leq\ & \frac{c_\beta 2^p |x_{k-N} - \bar{x}_{k-N}|^p}{N^q} + \\ & \beta\left(\frac{2\, 5^{1/a} \bar{c}_p^{1/a} |x_{k-N} - \bar{x}_{k-N}| g\left(k,l\right)^{1/a}}{\underline{c}_p^{1/a}} + \right.\\ & \frac{2\, 5^{1/a} \bar{c}_\Lambda^{1/a} |\alpha_{k-N} - \bar{\alpha}_{k-N}| g\left(k,l\right)^{1/a}}{\underline{c}_p^{1/a}} + \\ & \frac{2\, 5^{1/a} N^{1/a} \overline{\gamma}_w \left(\|\boldsymbol{w}\|_{[k-N,k-1]}\right)^{1/a} g\left(k,l\right)^{1/a}}{\underline{c}_p^{1/a}} + \\ & \frac{2\, 5^{1/a} N^{1/a} \overline{\gamma}_v \left(\|\boldsymbol{v}\|_{[k-N,k-1]}\right)^{1/a} g\left(k,l\right)^{1/a}}{\underline{c}_p^{1/a}} + \\ & \left.\frac{2\, 5^{1/a} N^{1/a} \overline{\gamma}_\alpha \left(\|\boldsymbol{w_\alpha}\|_{[k-N,k-1]}\right)^{1/a} g\left(k,l\right)^{1/a}}{\underline{c}_p^{1/a}}, N\right)\end{aligned} \tag{40}$$

Using inequalities (1)

$$\begin{aligned}\beta\left(|x_{k-N}-\hat{x}_{k-N|k}|,N\right) \leq\ & \frac{c_\beta 2^p |x_{k-N}-\bar{x}_{k-N}|^p}{N^q}+\\ & \beta\left(\frac{10\,5^{1/a}\overline{c}_p^{1/a}|x_{k-N}-\bar{x}_{k-N}|g\left(k,l\right)^{1/a}}{\underline{c}_p^{1/a}},N\right)+\\ & \beta\left(\frac{10\,5^{1/a}\overline{c}_\Lambda^{1/a}|\alpha_{k-N}-\bar{\alpha}_{k-N}|g\left(k,l\right)^{1/a}}{\underline{c}_p^{1/a}},N\right)+\\ & \beta\left(\frac{10\,5^{1/a}N^{1/a}\overline{\gamma}_w\left(\|\boldsymbol{w}\|_{[k-N,k-1]}\right)^{1/a}g\left(k,l\right)^{1/a}}{\underline{c}_p^{1/a}},N\right)+\\ & \beta\left(\frac{10\,5^{1/a}N^{1/a}\overline{\gamma}_v\left(\|\boldsymbol{v}\|_{[k-N,k-1]}\right)^{1/a}g\left(k,l\right)^{1/a}}{\underline{c}_p^{1/a}},N\right)+\\ & \beta\left(\frac{10\,5^{1/a}N^{1/a}\overline{\gamma}_\alpha\left(\|\boldsymbol{w_\alpha}\|_{[k-N,k-1]}\right)^{1/a}g\left(k,l\right)^{1/a}}{\underline{c}_p^{1/a}},N\right)+\end{aligned} \tag{41}$$

Now, by mean of Assumption 2

$$\begin{aligned}\beta\left(|x_{k-N}-\hat{x}_{k-N|k}|,N\right) \leq\ & \frac{c_\beta 2^p |x_{k-N}-\bar{x}_{k-N}|^p}{N^q}+\\ & \frac{c_\beta 10^p\,5^{p/a}\overline{c}_p^{p/a}|x_{k-N}-\bar{x}_{k-N}|^p g\left(k,l\right)^{p/a}}{N^q\,\underline{c}_p^{p/a}}+\\ & \frac{c_\beta 10^p\,5^{p/a}\overline{c}_\Lambda^{p/a}|\alpha_{k-N}-\bar{\alpha}_{k-N}|^p g\left(k,l\right)^{p/a}}{N^q\,\underline{c}_p^{p/a}}+\\ & \frac{c_\beta 10^p\,5^{p/a}N^{1/a}\overline{\gamma}_w\left(\|\boldsymbol{w}\|_{[k-N,k-1]}\right)^{1/a}g\left(k,l\right)^{1/a}}{N^q\,\underline{c}_p^{p/a}}+\\ & \frac{c_\beta 10^p\,5^{p/a}N^{p/a}\overline{\gamma}_v\left(\|\boldsymbol{v}\|_{[k-N,k-1]}\right)^{p/a}g\left(k,l\right)^{p/a}}{N^q\,\underline{c}_p^{p/a}}+\\ & \frac{c_\beta 10^p\,5^{p/a}N^{p/a}\overline{\gamma}_\alpha\left(\|\boldsymbol{w_\alpha}\|_{[k-N,k-1]}\right)^{p/a}g\left(k,l\right)^{p/a}}{N^q\,\underline{c}_p^{p/a}}\end{aligned} \tag{42}$$

Rearranging terms

$$\beta\left(|x_{k-N}-\hat{x}_{k-N|k}|,N\right) \leq \frac{|x_{k-N}-\bar{x}_{k-N}|^p}{N^q}\left(c_\beta 2^p+\frac{c_\beta 10^p\,5^{p/a}\overline{c}_p^{p/a}g\left(k,l\right)^{p/a}}{\underline{c}_p^{p/a}}\right)+ \frac{c_\beta 10^p\,5^{p/a}\overline{c}_\Lambda^{p/a}|\alpha_{k-N}-\bar{\alpha}_{k-N}|^p g\left(k,l\right)^{p/a}}{N^q\,\underline{c}_p^{p/a}}+ \frac{c_\beta 10^p\,5^{p/a}N^{1/a}\overline{\gamma}_w\left(\|\boldsymbol{w}\|_{[k-N,k-1]}\right)^{1/a}g\left(k,l\right)^{1/a}}{N^q\,\underline{c}_p^{p/a}}+ \frac{c_\beta 10^p\,5^{p/a}N^{p/a}\overline{\gamma}_v\left(\|\boldsymbol{v}\|_{[k-N,k-1]}\right)^{p/a}g\left(k,l\right)^{p/a}}{N^q\,\underline{c}_p^{p/a}}+ \frac{c_\beta 10^p\,5^{p/a}N^{p/a}\overline{\gamma}_\alpha\left(\|\boldsymbol{w_\alpha}\|_{[k-N,k-1]}\right)^{p/a}g\left(k,l\right)^{p/a}}{N^q\,\underline{c}_p^{p/a}} \tag{43}$$

Once we have found an upper bound for the first term of (38), we will follow a similar procedure to find a bound for the second and third terms. Using (35) for the $l-th$ iteration, we can write

$$|\hat{w}_{j|k}| \overset{(35)(1)}{\leq} \underline{\gamma}_w^{-1}\left(\frac{5\overline{c}_p|x_{k-N}-\bar{x}_{k-N}|^a g(k,l)}{N}\right)+\underline{\gamma}_w^{-1}\left(\frac{5\overline{c}_\Lambda|\alpha_{k-N}-\bar{x}_{k-N}|^a g(k,l)}{N}\right)+ \underline{\gamma}_w^{-1}\left(\frac{5g(k,l)\overline{\gamma}_w\left(\|\boldsymbol{w}\|_{[k-N,k]}\right)}{N}\right)+\underline{\gamma}_w^{-1}\left(\frac{5g(k,l)\overline{\gamma}_v\left(\|\boldsymbol{v}\|_{[k-N,k]}\right)}{N}\right)+ \underline{\gamma}_w^{-1}\left(\frac{5g(k,l)\overline{\gamma}_\alpha\left(\|\boldsymbol{w_\alpha}\|_{[k-N,k]}\right)}{N}\right) \tag{44}$$

Introducing this bound in the second term of Equation (38):

$$\begin{aligned}\gamma_1\left(\|\boldsymbol{w}-\hat{\boldsymbol{w}}\|_{[k-N,k]}\right) \leq&\gamma_1\left(\|\boldsymbol{w}\|_{[k-N,k]}+\|\hat{\boldsymbol{w}}\|_{[k-N,k]}\right)\\ \leq&\gamma_1\left(\|\boldsymbol{w}\|_{[k-N,k]}+\underline{\gamma}_w^{-1}\left(\frac{5\overline{c}_p|x_{k-N}-\bar{x}_{k-N}|^a g\left(k,l\right)}{N}\right)+\right.\\ &\underline{\gamma}_w^{-1}\left(\frac{5\overline{c}_\Lambda|\alpha_{k-N}-\bar{x}_{k-N}|^a g\left(k,l\right)}{N}\right)+\\ &\underline{\gamma}_w^{-1}\left(\frac{5g\left(k,l\right)\overline{\gamma}_w\left(\|\boldsymbol{w}\|_{[k-N,k]}\right)}{N}\right)+\\ &\underline{\gamma}_w^{-1}\left(\frac{5g\left(k,l\right)\overline{\gamma}_v\left(\|\boldsymbol{v}\|_{[k-N,k]}\right)}{N}\right)+\\ &\left.\underline{\gamma}_w^{-1}\left(\frac{5g\left(k,l\right)\overline{\gamma}_\alpha\left(\|\boldsymbol{w_\alpha}\|_{[k-N,k]}\right)}{N}\right)\right)\end{aligned} \tag{45}$$

Recalling Inequalities (1) we obtain the bound

$$\begin{aligned}\gamma_1\left(\|\boldsymbol{w}_j-\hat{\boldsymbol{w}}_{j|k}\|_{[k-N,k]}\right)\leq\ &\gamma_1\left(6\|\boldsymbol{w}\|_{[k-N,k]}\right)+\gamma_1\left(6\underline{\gamma}_w^{-1}\left(\frac{5\overline{c}_p|x_{k-N}-\bar{x}_{k-N}|^a g(k,l)}{N}\right)\right)+\\ &\gamma_1\left(6\underline{\gamma}_w^{-1}\left(\frac{5\overline{c}_\Lambda|\alpha_{k-N}-\bar{x}_{k-N}|^a g(k,l)}{N}\right)\right)+\\ &\gamma_1\left(6\underline{\gamma}_w^{-1}\left(\frac{5g(k,l)\overline{\gamma}_w\left(\|\boldsymbol{w}\|_{[k-N,k]}\right)}{N}\right)\right)+\\ &\gamma_1\left(6\underline{\gamma}_w^{-1}\left(\frac{5g(k,l)\overline{\gamma}_v\left(\|\boldsymbol{v}\|_{[k-N,k]}\right)}{N}\right)\right)+\\ &\gamma_1\left(6\underline{\gamma}_w^{-1}\left(\frac{5g(k,l)\overline{\gamma}_\alpha\left(\|\boldsymbol{w}_{\boldsymbol{\alpha}}\|_{[k-N,k]}\right)}{N}\right)\right).\end{aligned} \tag{46}$$

With the use of Assumption 3, the bound can be finally written as

$$\begin{aligned}\gamma_1\left(\|\boldsymbol{w}_j-\hat{\boldsymbol{w}}_{j|k}\|_{[k-N,k]}\right)\leq\ &\gamma_1\left(6\|\boldsymbol{w}\|_{[k-N,k]}\right)+\frac{c_1 5^{b_1}\overline{c}_p^{b_1}|x_{k-N}-\bar{x}_{k-N}|^{ab_1}g\left(k,l\right)^{b_1}}{N^{b_1}}+\\ &\frac{c_1 5^{b_1}\overline{c}_\Lambda^{b_1}|\alpha_{k-N}-\bar{\alpha}_{k-N}|^{ab_1}g\left(k,l\right)^{b_1}}{N^{b_1}}+\\ &c_1 5^{b_1}g\left(k,l\right)^{b_1}\overline{\gamma}_w\left(\|\boldsymbol{w}\|_{[k-N,k]}\right)^{b_1}+\\ &c_1 5^{b_1}g\left(k,l\right)^{b_1}\overline{\gamma}_v\left(\|\boldsymbol{v}\|_{[k-N,k]}\right)^{b_1}+\\ &c_1 5^{b_1}g\left(k,l\right)^{b_1}\overline{\gamma}_\alpha\left(\|\boldsymbol{w}_{\boldsymbol{\alpha}}\|_{[k-N,k]}\right)^{b_1}\end{aligned} \tag{47}$$

With a similar procedure, a bound for the third term of inequality (38) is found

$$\begin{aligned}\gamma_2\left(\|\boldsymbol{v}_j-\hat{\boldsymbol{v}}_{j|k}\|_{[k-N,k]}\right)\leq\ &\gamma_2\left(6\|\boldsymbol{v}\|_{[k-N,k]}\right)+\frac{c_2 5^{b_2}\overline{c}_p^{b_2}|x_{k-N}-\bar{x}_{k-N}|^{ab_2}g\left(k,l\right)^{b_2}}{N^{b_2}}+\\ &\frac{c_2 5^{b_2}\overline{c}_\Lambda^{b_2}|\alpha_{k-N}-\bar{\alpha}_{k-N}|^{ab_2}g\left(k,l\right)^{b_2}}{N^{b_2}}+\\ &c_2 5^{b_2}g\left(k,l\right)^{b_2}\overline{\gamma}_w\left(\|\boldsymbol{w}\|_{[k-N,k]}\right)^{b_2}+\\ &c_2 5^{b_2}g\left(k,l\right)^{b_2}\overline{\gamma}_v\left(\|\boldsymbol{v}\|_{[k-N,k]}\right)^{b_2}+\\ &c_2 5^{b_2}g\left(k,l\right)^{b_2}\overline{\gamma}_\alpha\left(\|\boldsymbol{w}_{\boldsymbol{\alpha}}\|_{[k-N,k]}\right)^{b_2}\end{aligned} \tag{48}$$

The estimation error given in Equation (49) can be bounded as

$$
\begin{aligned}
|x_k - \hat{x}_{k|k}| \leq \; & \frac{|x_{k-N} - \bar{x}_{k-N}|^p}{N^q}\left(c_\beta 2^p + \frac{c_\beta 10^p\, 5^{p/a} \overline{c}_p^{p/a} g\,(k,l)^{p/a}}{\underline{c}_p^{p/a}}\right) + \\
& \frac{c_\beta 10^p\, 5^{p/a} \overline{c}_\Lambda^{p/a} |\alpha_{k-N} - \bar{\alpha}_{k-N}|^p g\,(k,l)^{p/a}}{N^q\, \underline{c}_p^{p/a}} + \\
& \frac{c_\beta 10^p\, 5^{p/a} N^{1/a} \overline{\gamma}_w \left(\|\boldsymbol{w}\|_{[k-N,k-1]}\right)^{1/a} g\,(k,l)^{1/a}}{N^q\, \underline{c}_p^{p/a}} + \\
& \frac{c_\beta 10^p\, 5^{p/a} N^{p/a} \overline{\gamma}_v \left(\|\boldsymbol{v}\|_{[k-N,k-1]}\right)^{p/a} g\,(k,l)^{p/a}}{N^q\, \underline{c}_p^{p/a}} + \\
& \frac{c_\beta 10^p\, 5^{p/a} N^{p/a} \overline{\gamma}_\alpha \left(\|\boldsymbol{w_\alpha}\|_{[k-N,k-1]}\right)^{p/a} g\,(k,l)^{p/a}}{N^q\, \underline{c}_p^{p/a}} + \\
& \gamma_1\left(6\|\boldsymbol{w}\|_{[k-N,k]}\right) + \frac{c_1 5^{b_1} \overline{c}_p^{b_1} |x_{k-N} - \bar{x}_{k-N}|^{ab_1} g\,(k,l)^{b_1}}{N^{b_1}} + \\
& \frac{c_1 5^{b_1} \overline{c}_\Lambda^{b_1} |\alpha_{k-N} - \bar{\alpha}_{k-N}|^{ab_1} g\,(k,l)^{b_1}}{N^{b_1}} + \\
& c_1 5^{b_1} g\,(k,l)^{b_1}\, \overline{\gamma}_w \left(\|\boldsymbol{w}\|_{[k-N,k]}\right)^{b_1} + \\
& c_1 5^{b_1} g\,(k,l)^{b_1}\, \overline{\gamma}_v \left(\|\boldsymbol{v}\|_{[k-N,k]}\right)^{b_1} + \\
& c_1 5^{b_1} g\,(k,l)^{b_1}\, \overline{\gamma}_\alpha \left(\|\boldsymbol{w_\alpha}\|_{[k-N,k]}\right)^{b_1} + \\
& \gamma_2\left(6\|\boldsymbol{v}\|_{[k-N,k]}\right) + \frac{c_2 5^{b_2} \overline{c}_p^{b_2} |x_{k-N} - \bar{x}_{k-N}|^{ab_2} g\,(k,l)^{b_2}}{N^{b_2}} + \\
& \frac{c_2 5^{b_2} \overline{c}_\Lambda^{b_2} |\alpha_{k-N} - \bar{\alpha}_{k-N}|^{ab_2} g\,(k,l)^{b_2}}{N^{b_2}} + \\
& c_2 5^{b_2} g\,(k,l)^{b_2}\, \overline{\gamma}_w \left(\|\boldsymbol{w}\|_{[k-N,k]}\right)^{b_2} + \\
& c_2 5^{b_2} g\,(k,l)^{b_2}\, \overline{\gamma}_v \left(\|\boldsymbol{v}\|_{[k-N,k]}\right)^{b_2} + \\
& c_2 5^{b_2} g\,(k,l)^{b_2}\, \overline{\gamma}_\alpha \left(\|\boldsymbol{w_\alpha}\|_{[k-N,k]}\right)^{b_2}
\end{aligned} \tag{49}
$$

Since the vector $\alpha$ (and its estimated $\hat{\alpha}$) satisfies $\sum_{i=1}^{q} \alpha_i = 1$ and $\alpha_i \geq 0$, the maximal value of $|\alpha - \hat{\alpha}|$ is upper bounded by $\sqrt{2}$, i.e., $\max\{|\alpha - \hat{\alpha}|\} = \max\{|w_\alpha|\} \leq \sqrt{2}$. Defining the constants as follows

$$q \geq p/a,\; \zeta \coloneqq \max\{p, a\, b_1, a\, b_2\} \;\; \eta \coloneqq \min\{q, b_1, b_2\},$$

inequality (49) can be rewritten as follows

$$
\begin{aligned}
|x_k - \hat{x}_{k|k}| \leq \frac{|x_{k-N} - \bar{x}_{k-N}|^{\zeta}}{N^{\eta}} & \left( \left( 1 + \frac{5^{p+q}\overline{c}_p^q g\left(k,l\right)^q}{\underline{c}_p^{p/a}} \right) c_\beta 2^p + c_1 5^{b_1} \overline{c}_1^{b_1} g\left(k,l\right)^{b_1} + \right. \\
& \left. c_2 5^{b_2} \overline{c}_p^{b_2} g\left(k,l\right)^{b_2} \right) + \frac{|\alpha_{k-N} - \bar{\alpha}_{k-N}|^{\zeta}}{N^{\eta}} \left( \frac{c_\beta 10^p 5^q \overline{c}_\Lambda^q g\left(k,l\right)^q}{\overline{c}_p^{p/a}} + \right. \\
& \left. c_1 5^{b_1} \overline{\gamma}_\alpha \left(\sqrt{2}\right)^{b_1} + c_2 5^{b_2} \overline{\gamma}_\alpha \left(\sqrt{2}\right)^{b_2} \right) + g\left(k,l\right)^{\eta} \left( \frac{c_\beta 10^p 5^q \overline{\gamma}_\alpha \left(\sqrt{2}\right)^q}{\underline{c}_p^{p/a}} + \right. \\
& \left. c_1 5^{b_1} \overline{\gamma}_\alpha \left(\sqrt{2}\right)^{b_1} + c_2 5^{b_2} \overline{\gamma}_\alpha \left(\sqrt{2}\right)^{b_2} \right) + \frac{c_\beta 10^p 5^q g(k,l)^q \overline{\gamma}_w \left(\|\boldsymbol{w}\|_{[k-N,k]}\right)^q}{\underline{c}_p^{p/a}} + \\
& \gamma_1 \left(6\|\boldsymbol{w}\|_{[k-N,k]}\right) + c_1 5^{b_1} g\left(k,l\right)^{b_1} \overline{\gamma}_w \left(\|\boldsymbol{w}\|_{[k-N,k]}\right)^{b_1} + \\
& c_2 5^{b_2} g\left(k,l\right)^{b_2} \overline{\gamma}_w \left(\|\boldsymbol{w}\|_{[k-N,k]}\right)^{b_2} + \frac{c_\beta 10^p 5^q g(k,l)^q \overline{\gamma}_v \left(\|\boldsymbol{v}\|_{[k-N,k]}\right)^q}{\underline{c}_p^{p/a}} + \\
& \gamma_2 \left(6\|\boldsymbol{v}\|_{[k-N,k]}\right) + c_1 5^{b_1} g\left(k,l\right)^{b_1} \overline{\gamma}_v \left(\|\boldsymbol{v}\|_{[k-N,k]}\right)^{b_1} + \\
& c_2 5^{b_2} g\left(k,l\right)^{b_2} \overline{\gamma}_v \left(\|\boldsymbol{v}\|_{[k-N,k]}\right)^{b_2}.
\end{aligned}
\tag{50}
$$

Noting again that $|\alpha_{k-N} - \bar{\alpha}_{k-N}| \leq \sqrt{2}$ and defining the functions and constants as follows

$$k_1 := c_\beta 2^p \tag{51}$$

$$k_2 := \left( 5^{p+q} \left( \frac{\overline{c}_p}{\underline{c}_p} \right)^q c_\beta 2^p + c_1 5^{b_1} \overline{c}_p^{b_1} + c_2 5^{b_2} \overline{c}_p^{b_2} \right), \tag{52}$$

$$K := \frac{c_\beta 10^p 5^q}{\underline{c}_p^{p/a}} + c_1 5^{b_1} + c_2 5^{b_2} \tag{53}$$

$$\psi_{w1} := \gamma_1 \left(6\|\boldsymbol{w}\|_{[k-N,k]}\right), \tag{54}$$

$$\psi_{w2} := \frac{c_\beta 10^p 5^q \overline{\gamma}_w \left(\|\boldsymbol{w}\|_{[k-N,k]}\right)^q}{\underline{c}_p^{p/a}} + \left(c_1 5^{b_1} + c_2 5^{b_2}\right) \overline{\gamma}_w \left(\|\boldsymbol{w}\|_{[k-N,k]}\right)^{b_1} \tag{55}$$

$$\psi_{v1} := \gamma_2 \left(6\|\boldsymbol{v}\|_{[k-N,k]}\right), \tag{56}$$

$$\psi_{w2} := \frac{c_\beta 10^p 5^q \overline{\gamma}_v \left(\|\boldsymbol{v}\|_{[k-N,k]}\right)^q}{\underline{c}_p^{p/a}} + \left(c_1 5^{b_1} + c_2 5^{b_2}\right) \overline{\gamma}_v \left(\|\boldsymbol{v}\|_{[k-N,k]}\right)^{b_1} \tag{57}$$

(58)

the bound of the estimation error can be rewritten as follows

$$
\begin{aligned}
|x_k - \hat{x}_{k|k}| \leq \frac{|x_{k-N} - \bar{x}_{k-N}|^{\zeta}}{N^{\eta}} & \left(k_1 + g\left(k,l\right)^{\eta} k_2\right) + \psi_{w1} + g\left(k,l\right) \psi_{w2} + \\
& \psi_{v1} + g\left(k,l\right) \psi_{v2} + g\left(k,l\right) \left( \overline{\gamma}_\alpha \left(\sqrt{2}\right)^{\zeta} + \frac{\overline{c}_\Lambda 2^{\zeta/2}}{N^{\eta}} \right) K
\end{aligned}
\tag{59}
$$

Defining others constants and functions

$$\bar{k}_\beta(l) := k_1 + g(k,l)\, k_2 \tag{60}$$
$$\Phi_w(w,l) := \psi_{w1} + g(k,l)^\eta\, \psi_{w2} \tag{61}$$
$$\Phi_v(v,l) := \psi_{v1} + g(k,l)^\eta\, \psi_{v2} \tag{62}$$
$$\Phi_\alpha(l,N) := g(k,l)\left(\overline{\gamma}_\alpha\left(\sqrt{2}\right)^\zeta + \frac{\overline{c}_\Lambda^\zeta\, 2^{\zeta/2}}{N^\eta}\right) K \tag{63}$$
$$\bar{\beta}(r,s) := \frac{\bar{k}_\beta r^p}{s^q} \tag{64}$$

with $p = \zeta$, $q = \eta$, $\bar{\beta}(r,s) \in \mathscr{K}\mathscr{L}$, one can write the estimation error as follows

$$|x_k - \hat{x}_{k|k}| \leq \bar{\beta}(|x_{k-N} - \bar{x}_{k-N}|, N) + \Phi_w\left(\|\boldsymbol{w}\|_{[k-N,k]}, l\right) + \Phi_v\left(\|\boldsymbol{v}\|_{[k-N,k]}, l\right) + \Phi_\alpha(l,N) \tag{65}$$

The reader can verify that the same result is obtained for $k \in \mathbb{Z}_{[1,N-1]}$. To guarantee the validity of previous results to the entire time horizon the definition of $\beta(r,s)$ must be extended to $s = 0$ . Because of $\bar{\beta}(r,s) \in \mathscr{K}\mathscr{L}$, $\bar{\beta}(r,0) \in \mathscr{K}\mathscr{L}$ and $\bar{\beta}(r,0) \geq \bar{\beta}(r,k)$, $\forall k \in \mathbb{Z}_{\geq 1}$, it is sufficient to define $\bar{\beta}(r,0) := k_\beta\ \bar{\beta}(r,1)$ for some $k_\beta \in \mathbb{R}_{>1}$ to extend the definition of these function for all $k \in \mathbb{Z}_{[0,N]}$.

Let us select some $\epsilon \in \mathbb{R}_{>0}$ and

$$r_{\max} := \Big\{\bar{\beta}(e_{\max},0) + \Phi_w(w_{\max},1) + \Phi_v(v_{\max},1) + \Phi_\alpha(1,\mathscr{N}), \\ (1+\epsilon)\left(\Phi_w(w_{\max},1) + \Phi_v(v_{\max},1) + \Phi_\alpha(1,\mathscr{N})\right)\Big\}$$

Let us define $\mathscr{N}$ as

$$\left(\frac{2(1+\epsilon)(k_1 + k_2 g(k,l)^\eta)\, e_{\max}^\zeta}{r_{\max}}\right)^{1/\eta} \leq \mathscr{N} \tag{66}$$

**Remark 2** *Note that due to the model uncertainty, the horizon must be enlarged.*

Adopting an estimator with a window length greater or equal to $\mathscr{N}$, one will have

$$\bar{\beta}(r,\ N) \leq \frac{r}{2}, \tag{67}$$

the effects of the initial conditions will vanish. As $k \to \infty$, the estimation error will entry to the bounded set $\mathscr{X}(w,v) \subset \mathscr{X}$ defined by the noises of the system and the uncertainty

$$\mathscr{X}(w,v,l) := \{|x_k - \hat{x}_{k|k}| \leq (1+\epsilon)\left(\Phi_w(\|\boldsymbol{w}\|,1) + \Phi_v(\|\boldsymbol{v}\|,1) + \Phi_\alpha(1,\mathscr{N})\right)\}. \tag{68}$$

This set define the minimum size region of error space $\mathscr{X}$ that the error can achieve by removing the effect of errors in initial conditions ($e_{max}$). Equation (67) establish a trade off between speed of convergence and window length, which is related with the size of $\mathscr{X}(w, v, l)$.

For any MHE with adaptive arrival cost and window length $N \geq \mathscr{N}$ two situations can be considered

- The estimator has removed the effects of $x_0$ on $\hat{x}_{k|k}$ such that $|x_k - \hat{x}_{k|k}| \in \mathscr{X}(w, v, l)$, and
- The estimator has not removed the effects of $x_0$ on $\hat{x}_{k|k}$ such that $|x_k - \hat{x}_{k|k}| \notin \mathscr{X}(w, v, l)$,

Let us assume that the estimation error is

$$|x_k - \hat{x}_{k|k}| \leq 2(1+\epsilon)\left(\Phi_w(w_{\max}, 1) + \Phi_v(v_{\max}, 1) + \Phi_\alpha(1, \mathscr{N})\right),$$

the estimation error will be given by

$$\begin{aligned}|x_{k+N} - \hat{x}_{k+N|k+N}| \leq& \bar{\beta}\left(|x_k - \hat{x}_{k|k+N}|, \mathscr{N}\right) + \Phi_w(w_{\max}, 1) + \\ & \Phi_v(v_{\max}, 1) + \Phi_\alpha(1, \mathscr{N}), \\ \leq& \frac{|x_k - \hat{x}_{k|k}|}{2(1+\epsilon)} + \Phi_w(w_{\max}, 1) + \Phi_v(v_{\max}, 1) + \Phi_\alpha(1, \mathscr{N}) \\ \leq& 2\left(\Phi_w(w_{\max}, 1) + \Phi_v(v_{\max}, 1) + \Phi_\alpha(1, \mathscr{N})\right) \\ \leq& 2(1+\epsilon)\left(\Phi_w(w_{\max}, 1) + \Phi_v(v_{\max}, 1) + \Phi_\alpha(1, \mathscr{N})\right)\end{aligned} \tag{69}$$

Therefore, when

$$|x_k - \hat{x}_{k|k}| \leq 2(1+\epsilon)\left(\Phi_w(w_{\max}, 1) + \Phi_v(v_{\max}, 1) + \overline{K}(1)\right),$$

the error will no become larger. Assuming now

$$r_{\max} \geq |x_k - \hat{x}_{k|k}| > 2(1+\epsilon)\left(\Phi_w(w_{\max}) + \Phi_v(v_{\max}) + \Phi_\alpha(1, \mathscr{N})\right)$$

the estimation error is given by

$$\begin{aligned}|x_{k+N} - \hat{x}_{k+N|k+N}| \leq& \bar{\beta}\left(|x_k - \hat{x}_{k|k+N}|, \mathscr{N}\right) + \Phi_w(w_{\max}, 1) + \\ & \Phi_v(v_{\max}, 1) + \Phi_\alpha(1, \mathscr{N}) \\ \leq& \tfrac{|x_k - \hat{x}_{k|k+N}|}{2(1+\epsilon)} + \Phi_w(w_{\max}, 1) + \Phi_v(v_{\max}, 1) + \Phi_\alpha(1, \mathscr{N}) \\ \leq& \tfrac{|x_k - \hat{x}_{k|k+N}|}{2(1+\epsilon)} + \tfrac{|x_k - \hat{x}_{k|k+N}|}{2(1+\epsilon)} \\ \leq& |x_k - \hat{x}_{k|k+N}|\left(\tfrac{1}{1+\epsilon}\right)\end{aligned}$$

with $\xi := \left(\frac{1}{1+\epsilon}\right) < 1$, since $\epsilon > 0$

In the latter case, the estimator error behaves contractively. By mean of some definitions, i.e., $i \coloneqq \lfloor \frac{k}{\mathscr{N}} \rfloor$, $j \coloneqq t \mod \mathscr{N}$, time $k$ can be expressed as $k = i\mathscr{N} + j$. For $N \geq \mathscr{N}$, the $\mathscr{KL}$ functions $\bar{\beta}(r, s)$ is decreasing every $\mathscr{N}$ samples. Writing the estimation error with this notation for time $k$

$$\begin{aligned} |x_k - \hat{x}_{k|k}| &\leq \max \left\{ \xi^i |x_j - \hat{x}_{j|k}|, 2(1+\epsilon)\left(\Phi_w(w_{\max}, 1) + \Phi_v(v_{\max}, 1) + \Phi_\alpha(1, \mathscr{N})\right) \right\} \\ &\leq \bar{\beta}(|x_0 - \bar{x}_0|, j)\, \xi^i + 2(1+\mu)\left(\Phi_w(w_{\max}, 1) + \Phi_v(v_{\max}, 1) + \Phi_\alpha(1, \mathscr{N})\right) \end{aligned} \tag{70}$$

Defining

$$\begin{aligned} \overline{\Phi}(|x_0 - \bar{x}_0|) &\coloneqq \xi^i \bar{\beta}(|x_0 - \bar{x}_0|, j), \\ \overline{\Phi}_w(w_{\max}) &\coloneqq 2(1+\mu)\Phi_w(w_{\max}, 1), \\ \overline{\Phi}_v(v_{\max}) &\coloneqq 2(1+\mu)\Phi_v(v_{\max}, 1), \\ \overline{\Phi}_\alpha(l_{\min}, \mathscr{N}) &\coloneqq 2(1+\mu)\Phi_\alpha(1, \mathscr{N}). \end{aligned} \tag{71}$$

Finally we can write

$$|x_k - \hat{x}_{k|k}| \leq \overline{\Phi}(|x_0 - \bar{x}_0|, k) + \overline{\Phi}_w(w_{\max}) + \overline{\Phi}_v(v_{\max}) + \overline{\Phi}_\alpha(l_{\min}, \mathscr{N}) \tag{72}$$

Taking $w_{\max}$ from $\|\boldsymbol{w}\|_{[0,k]}$ instead $\|\boldsymbol{w}\|_{[k-N,k]}$ and $v_{\max}$ from $\|\boldsymbol{v}\|_{[0,k]}$ instead $\|\boldsymbol{v}\|_{[k-N,k]}$, and noting that Equation (65) still being valid, the robust regional practical stability is proved.

On the other hand, the convergence of the estimator to the true state in the case of decaying disturbances can be established. Assuming $\lim_{k\to\infty} w_k = 0$ and $\lim_{k\to\infty} v_k = 0$, given Equation (65), one can choose some $k \geq K_1$ for which $\max\{w_{\max}, v_{\max}\} \leq \min\left\{\Phi_w^{-1}\left(\frac{\varepsilon}{4}\right), \Phi_v^{-1}\left(\frac{\varepsilon}{4}\right)\right\}$. At the same time, one can choose some large enough value of $l$ such that $\overline{\Phi}_\alpha(l, \mathscr{N}) \leq \frac{\varepsilon}{4}$. Note that according to Equation (32), the value of $l$ will be getting smaller as $\|\boldsymbol{w}\| \to 0$ and $\|\boldsymbol{v}\| \to 0$. Recalling that $\overline{\Phi}(\cdot) \in \mathscr{KL}$, there will exist some $k \geq K_2$ such that $\overline{\Phi}(|x_{k-K_2} - \bar{x}_{k-K_2}|, k) \leq \frac{\varepsilon}{4}$. Under these conditions, there exists some time $k \geq \max\{K_1, K_2\} + \mathscr{N}$ such that

$$\begin{aligned} &|x_k - \hat{x}_{k|k}| \leq \overline{\Phi}(|x_{k-\mathscr{N}} - \bar{x}_{k-\mathscr{N}}|, k) + \overline{\Phi}_w(w_{\max}) + \overline{\Phi}_v(v_{\max}) + \overline{\Phi}_\alpha(l, \mathscr{N}) \\ &|x_k - \hat{x}_{k|k}| \leq \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} \\ &|x_k - \hat{x}_{k|k}| \leq \varepsilon \end{aligned} \tag{73}$$

Since one can choose any value of $\varepsilon$, $\lim_{k\to\infty} |x_k - \hat{x}_{k|k}| = 0$ can be guaranteed when $\lim_{k\to\infty} w_k = 0$ and $\lim_{k\to\infty} v_k = 0$ as claimed.

□

**Remark 3** *As expected, the model uncertainty deteriorates state estimation.*

*However, for a large enough value of* $l$*, this effect can be mitigated. Moreover, when* $\lim_{l\to\infty}\Phi_\alpha\left(l,\mathscr{N}\right)=0$

## 4 Simulation and results

The following examples will be used to illustrate the results presented in the previous sections and to evaluate and compare the performance of the proposed estimator with others from the state of the art. The process disturbance is white Gaussian noise acting as an additive exogenous input to the system.

### *4.1 Unknown linear system*

Let us considers the linear system

$$\begin{aligned} x_{k+1} &= A_p x_k + w_k \\ y_k &= C_p x_k + v_k \end{aligned}$$

whose matrices are unknown an they have the following structure

$$A_p = \begin{bmatrix} 0 & a_1 \\ 1 & a_2 \end{bmatrix}, C_p = \begin{bmatrix} c_1 & c_2 \end{bmatrix} \tag{74}$$

The system is affected with additive process and measurement noise $w$ and $v$ drawn from normal distributions with zero mean and covariance $Q_w = S_w^2 I_2$ and $R_v = S_v^2$, respectively. The polytope is defined using three $LTI$ models

$$\begin{aligned} A_1 &= \begin{bmatrix} 0 & 0.72 \\ 1 & 0.28 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & -0.59 \\ 1 & 1.57 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & -0.35 \\ 1 & 1.26 \end{bmatrix}, \\ C_1 &= \begin{bmatrix} -1.46 & -1.29 \end{bmatrix}, C_2 = \begin{bmatrix} -4.84 & -2.90 \end{bmatrix}, C_3 = \begin{bmatrix} -0.09 & -0.03 \end{bmatrix}, \end{aligned} \tag{75}$$

such that the system belongs to it. The matrices of the system $(A_S, C_S)$ and its model $(A_M, C_M)$ were generated as a convex combination of the polytope with mixing parameters $\alpha_{\mathbf{S}}$ $(\alpha_{S,1} = 0.22, \alpha_{S,2} = 0.76, \alpha_{S,3} = 0.02)$ and $\alpha_{\mathbf{M}}$ $(\alpha_{M,1} = 0.41, \alpha_{M,2} = 0.22, \alpha_{M,3} = 0.37)$, respectively.

The stage cost of the receding horizon estimators is chosen as $\ell(w,v) = w^T Q^{-1} w + v^T R^{-1} v$ with $R^{-1} = 5$ and $Q^{-1} = diag\,(0.1, 0.1)$. The proposed moving horizon estimator $(MHE_A)$ the prior weighting matrix is given by $\Gamma_{k-N}\left(\chi\right) = (\chi - \hat{x}(0|k))^T P_{k|k}^{-1}(\chi - \hat{x}(0|k))$, where $P_0^{-1} = 0.1\boldsymbol{I}_2$ and $P_{k|k}$ is updated using equations (8) and (9) with $\sigma = 1e^{-4}$ and $c = 5$. The robust

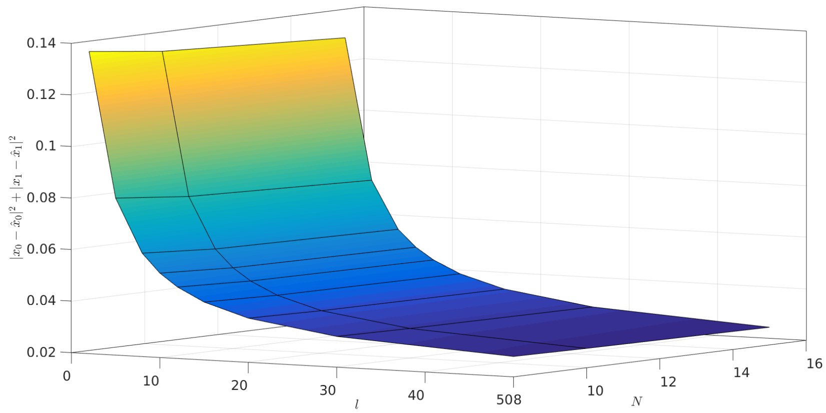


Fig. 1. Estimation error for different values of $l$ and $N$.

moving horizon estimator ($MHE_R$) implements the algorithm proposed by Müller (2017) with the nominal model $(A_M, C_M)$, the prior weight given by $\Gamma(\chi) = L(\chi - \hat{x}(0|k))^T(\chi - \hat{x}(0|k))$ and parameters $\delta = 1$, $\delta_1 = \kappa^N$ ($\kappa = 0.89$) and $\delta_2 = 1/N$ (see equation (3) of Müller (2017)). The full information (*FIE*, see Ji et al. (2016)) is configured with the true model and the same parameters used by the $MHE_R$ with $\delta = 1$, $\delta_1 = \kappa^k$ and $\delta_2 = 1/k$ with the system matrices $(A_S, C_S)$ The robust Kalman filter ($KF_R$) was designed following the design procedure proposed by Zhu et al. (2002) using the nominal model $(A_M, C_M)$ and computing the bounds from the models of the polytope. The Kalman filter ($KF$) was designed using the matrices of the system $(A_S, C_S)$

Table 1
Averaged MSE for $S_w = 0.1, S_v = 0.05, N = 8$.

| | $KF_R$ | $KF$ | $FIE$ | $MHE_R$ | $MHE_A$ |
|---|---|---|---|---|---|
| $x_0$ | 0.78467 | 0.030851 | 0.0054 | 0.2662 | 0.0212 |
| $x_1$ | 1.9946 | 0.069122 | 0.0039 | 0.4675 | 0.0389 |

Table 1 shows the mean square estimation error ($MSE$) of each estimator averaged over 100 trials for $S_w = 0.1$, $S_v = 0.05$ and $N = 8$ for all receding horizon estimators ($FIE$, $MHE_R$ and $MHE_A$). It can be seen that the proposed estimator average mean square estimation error is smaller than the Kalman filters ($KF$ and $KF_R$) and $MHE_R$. Only the $FIE$ provides better performance than the proposed algorithm. The performance difference between the estimators that employ the nominal model ($KF_R$ and $MHE_R$) is due to the adaptation capabilities of $MHE_A$ that allows to reduce the uncertainty of the estimator model. The main performance difference between $MHE_A$ and $FIE$ estimators is due to the model employed by estimator and the amount of information employed to estimate $\hat{x}_{k|k}$. While the $FIE$ estimator use of the exact model and all the system output available until $k$, the $MHE_A$ identifies the model in the initial samples and only use the last $N$ system outputs.

Figure 1 shows the behaviour of the estimation error as a function of $l$ and $N$. This figure shows that the main factor in the reduction of the estimation error is the number of iterations $l$ used to update the estimates. It can be also seen that there a significant improvement in the initial iterations ($l < 10$), then after iterations there is no significant improvement in the estimation error. It is worth nothing that the error is decreasing with the iterations as it was shown in Section 4.

Figure 2 shows the time evolution of the estimated vector of mixing parameters $\hat{\alpha}$ for different values of process noise variance. The true values are representing as continuous line. When $S_w$ is smaller than the value of states ($S_w \leq 0.5$), the mixing parameters $\hat{\alpha}$ converge quickly to the true value or remain closer to it.

### *4.2 Example 2: Nonlinear time-varying system*

As a second example, we consider a second order time–varying nonlinear system whose dynamic is given by behavior

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} x_2\left(k\right) \\ p_1\left(k\right) x_1\left(k\right) + \sin(p_2\, x_2\left(k\right)) \end{bmatrix} + w_k \\ y_k &= C x_k + v_k \end{aligned} \tag{76}$$

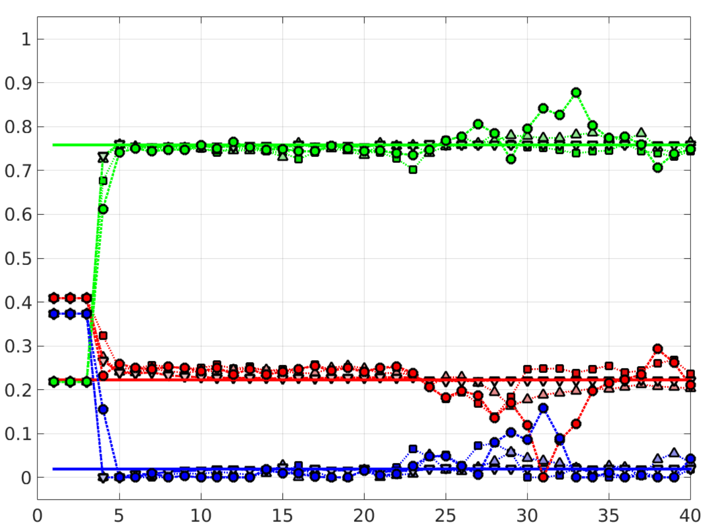


Fig. 2. Estimated mixing parameters $\hat{\alpha}$ for $N = 8$, $l =$, and noises variance ($\triangledown$ $S_w = 0.0$, $\square$ $S_w = 0.1$, $\triangle$ $S_w = 0.5$ and $\circ$ $S_w = 1.0$)

with the parameters $p_1(k)$ and $p_2(k)$ given by

$$\begin{aligned} p_1(k+1) &= 0.01\, p_1(k) \sin\left(\tfrac{5\pi k}{N}\right) && \forall k \in \mathbb{Z}_{[1,3N/4]}, \\ p_1(k+1) &= p_1(k) && \forall k \in \mathbb{Z}_{>3N/4}, \\ p_2 &= 0.05 \end{aligned} \quad (77)$$

The polytope was designed to guarantee that the nonlinear system always remains inside it. The polytope is defined using three $LTI$ models

$$A_1 = \begin{bmatrix} 0 & 1.30 \\ 1 & -1.52 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & -2.44 \\ 1 & 0.66 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 1.31 \\ 1 & 2.81 \end{bmatrix}, C = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}. \quad (78)$$

The stage cost for all receding horizon estimators is chosen as $\ell(w, v) = w^T Q^{-1} w + v^T R^{-1} v$ with $R^{-1} = 5e^2$ and $Q^{-1} = diag\,(1e^3, 5e^3)$. The proposed moving horizon estimator ($MHE_A$) the prior weighting matrix is given by $\Gamma_{k-N}(\chi) = (\chi - \hat{x}(0|k))^T P_{k|k}^{-1}(\chi - \hat{x}(0|k))$, where $P_0^{-1} = 0.1\boldsymbol{I}_2$ and $P_{k|k}$ is updated using equations (8) and (9) with $\sigma = 1$ and $c = 1$. The robust moving horizon estimator ($MHE_R$) implements the algorithm proposed by Müller (2017) with the nominal model, the prior weight given by $\Gamma(\chi) = L(\chi - \hat{x}(0|k))^T(\chi - \hat{x}(0|k))$ and parameters $\delta = 1$, $\delta_1 = \kappa^N$ ($\kappa = 0.89$) and $\delta_2 = 1/N$ (see equation (3) of Müller (2017)). The full information ($FIE$, see Ji et al. (2016)) is configured with the linearized model updated at each sampling time and the same parameters used by the $MHE_{ROB}$ with $\delta = 1$, $\delta_1 = \kappa^k$ and $\delta_2 = 1/k$. The robust Kalman filter was designed following the design procedure proposed by Zhu et al. (2002) using the nominal model and computing the bounds from the models of the polytope. The guess for the initial condition is $\bar{x}_0 = [0, 0]^T$, whereas $x_0 = [0.5, 0.3]^T$.

Table 2
Averaged MSE for $Sw = 0.1$, $Sv = 0.05$ and $N = 8$.

| | $EKF_R$ | $FIE$ | $MHE_R$ | $MHE_A$ |
|---|---|---|---|---|
| $x_0$ | 0.40879 | 0.32324 | 0.31541 | 0.018498 |
| $x_1$ | 0.4297 | 0.30082 | 1.0106 | 0.064984 |

Table 2 shows the mean square estimation error ($MSE$) of each estimator averaged over 100 trials for $S_w = 0.1$, $S_v = 0.05$ and $N = 8$ for all receding horizon estimators ($FIE$, $MHE_R$ and $MHE_A$). It can be seen that the average mean square estimation error of $MHE_A$ is smaller than the other estimators ($EKF_R$, $MHE_R$ and $FIE$). The performance difference between the estimators that employ the nominal model ($EKF_R$ and $MHE_R$) is mainly due to the adaptation capabilities of $MHE_A$. The main performance difference between $MHE_A$ and $FIE$ estimators is due to the FIE attempts to reconstruct the state trajectory of a nonlinear time-varying system with a $LTI$ system.

Figure 3 shows the behaviour of the estimation error as a function of $l$ and $N$. This figure shows that the main factor in the reduction of the estimation error is the number of iterations $l$ used to update the estimates. It can be also seen that there a significant improvement in the initial iterations ($l < 15$), then after iterations there is no significant improvement in the estimation error. It can also see how the estimation error increases for higher values of $N$. This behavior is due to the estimator use only one model along the entire estimation horizon, whereas the nonlinear system is changing its parameters every sample. The $EKF$ aim to improve the estimation error in comparison with the $MHE_A$. This improvement is due to the $EKF$ update the model at every sampling time with the true model. Besides, the $MHE_A$ use the same model along the entire horizon estimation to estimate the optimal state trajectory of the nonlinear system.

Figure 4 shows the time evolution of the estimated parameters $\hat{p}_1$ and $\hat{p}_2$ for

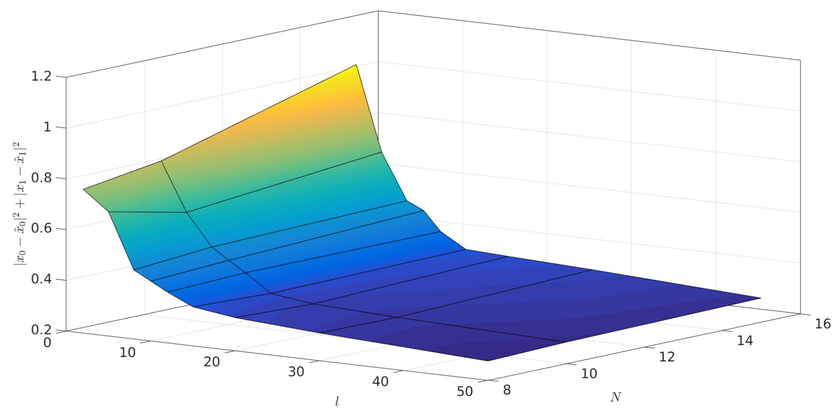


Fig. 3. Performance of the proposed algorithm for different values of $l$ and $N$.

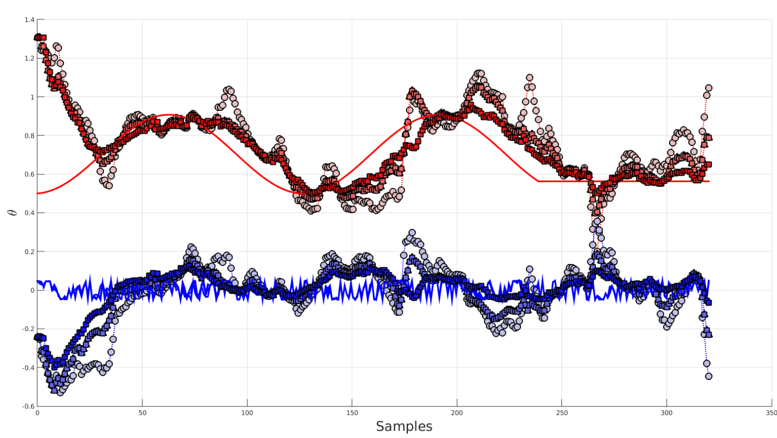


Fig. 4. Estimation of the parameters of the nonlinear time-varying system ( $N = 8$, and noises □=0.1, △=0.25 and ○=0.5).

different values of process noise variance. The true values are representing as continuous line. When $S_w$ is smaller than the value of states ($S_w \leq 0.25$), the parameters $\hat{p}$ converge quickly to the true value or remain closer to it. However,

## 5 Conclusions

In this paper we introduce an adaptive polytopic observer for nonlinear systems under bounded disturbances based on moving horizon estimator and dual estimation techniques and proved their stability properties. In a first stage we proved the stability of the dual estimation iteration. Then, in a second stage we established robust asymptotic stability for the adaptive moving horizon estimator. It was also shown that the estimation error converges to zero in case that disturbances converge to zero.

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.

## Acknowledgements

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

## References

Ames, W. F. & Pachpatte, B. (1997), *Inequalities for differential and integral equations*, Vol. 197, Elsevier.

Angelis, G. Z. (2003), 'System analysis, modelling and control with polytopic linear models.'.

Aouaouda, S., Chadli, M., Cocquempot, V. & Tarek Khadir, M. (2013), 'Multi-objective h-/ h fault detection observer design for takagi–sugeno fuzzy systems with unmeasurable premise variables: descriptor approach', *International Journal of Adaptive Control and Signal Processing* **27**(12), 1031–1047.

Apkarian, P., Gahinet, P. & Becker, G. (1995), 'Self-scheduled h control of linear parameter-varying systems: a design example', *Automatica* **31**(9), 1251–1261.

Arasaratnam, I. & Haykin, S. (2009), 'Cubature kalman filters', *IEEE Transactions on automatic control* **54**(6), 1254–1269.

Beneš, V. (1981), 'Exact finite-dimensional filters for certain diffusions with nonlinear drift', *Stochastics: An International Journal of Probability and Stochastic Processes* **5**(1-2), 65–92.

Bezzaoucha, S., Marx, B., Maquin, D. & Ragot, J. (2013), 'Nonlinear joint state and parameter estimation: Application to a wastewater treatment plant', *Control Engineering Practice* **21**(10), 1377–1385.

Bezzaoucha, S., Voos, H. & Darouach, M. (2018), 'A new polytopic approach for the unknown input functional observer design', *International Journal of Control* **91**(3), 658–677.

Crisan, D., Gaines, J. & Lyons, T. (1998), 'Convergence of a branching particle method to the solution of the zakai equation', *SIAM Journal on Applied Mathematics* **58**(5), 1568–1590.

Doucet, A., De Freitas, N. & Gordon, N. (2001), An introduction to sequential monte carlo methods, *in* 'Sequential Monte Carlo methods in practice', Springer, pp. 3–14.

Gao, N., Darouach, M., Voos, H. & Alma, M. (2016), 'New unified h dynamic observer design for linear systems with unknown inputs', *Automatica* **65**, 43–52.

Holte, J. M. (2009), Discrete gronwall lemma and applications, *in* 'MAA-NCS meeting at the University of North Dakota', Vol. 24, pp. 1–7.

Ichalal, D., Mammar, S. & Ragot, J. (2016), 'Auxiliary dynamics for observer design of nonlinear ts systems with unmeasurable premise variables', *IFAC-PapersOnLine* **49**(5), 1–6.

Ji, L., Rawlings, J. B., Hu, W., Wynn, A. & Diehl, M. (2016), 'Robust stability of moving horizon estimation under bounded disturbances', *IEEE Transactions on Automatic Control* **61**(11), 3509–3514.

Kushner, H. J. (1977), *Probability methods for approximations in stochastic control and for elliptic equations*, Vol. 129, Academic Press.

Leith, D. J. & Leithead, W. E. (2000), 'Survey of gain-scheduling analysis and design', *International journal of control* **73**(11), 1001–1025.

Müller, M. A. (2017), 'Nonlinear moving horizon estimation in the presence of bounded disturbances', *Automatica* **79**, 306–314.

Osorio-Gordillo, G. L., Darouach, M., Astorga-Zaragoza, C. M. & Boutat-Baddas, L. (2016), 'New dynamical observers design for linear descriptor systems', *IET Control Theory & Applications* **10**(17), 2223–2232.

Patwardhan, S. C., Narasimhan, S., Jagadeesan, P., Gopaluni, B. & Shah, S. L. (2012), 'Nonlinear bayesian state estimation: A review of recent developments', *Control Engineering Practice* **20**(10), 933–953.

Sánchez, G., Murillo, M. & Giovanini, L. (2017), 'Adaptive arrival cost update for improving moving horizon estimation performance', *ISA transactions* **68**, 54–62.

Shamma, J. S. & Athans, M. (1991), 'Guaranteed properties of gain scheduled control for linear parameter-varying plants', *Automatica* **27**(3), 559–564.

Shamma, J. S. & Cloutier, J. R. (1993), 'Gain-scheduled missile autopilot design using linear parameter varying transformations', *Journal of guidance, Control, and dynamics* **16**(2), 256–263.

Sontag, E. D. & Wang, Y. (1997), 'Output-to-state stability and detectability of nonlinear systems', *Systems & Control Letters* **29**(5), 279–290.

Theilliol, D. & Aberkane, S. (2011), 'Design of lpv observers with immeasurable gain scheduling variable under sensor faults', *IFAC Proceedings Volumes* **44**(1), 7613–7618.

Tóth, R., Abbas, H. S. & Werner, H. (2011), 'On the state-space realization of lpv input-output models: Practical approaches', *IEEE transactions on control systems technology* **20**(1), 139–153.

Zhu, X., Soh, Y. C. & Xie, L. (2002), 'Design and analysis of discrete-time robust kalman filters', *Automatica* **38**(6), 1069–1077.