Nonlinear Moving Horizon Estimation: Adaptive Arrival Cost with Prescribed Conditioning Number

Nestor Deniz*†, Guido Sanchez*, Marina Murillo* and Leonardo Giovanini*
*Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, sinc(i), UNL-CONICET,
Ciudad Universitaria UNL, 4to piso FICH (S3000), Santa Fe, Argentina.
†Email: ndeniz@sinc.unl.edu.ar

***Abstract*—In this paper, we propose a tuning mechanism for the arrival cost (AC) of nonlinear moving horizon estimators (NMHE) with quadratic stage cost (SC) that are robustly global asymptotically stable (RGAS). We consider general detectable nonlinear systems subject to bounded disturbances that are a priori unknown. Based on the robust stability results for NMHE, which state an optimal relation between the eigenvalues of the AC and SC weighting matrices in the sense of minimising the estimation error, we propose a mechanism which preserve the optimal conditioning number of the AC in every sampling time, maintaining low at the same time the computational burden. The performance of the proposed mechanism is illustrated through simulation studies.**

***Index Terms*—Nonlinear moving horizon estimation, Adaptive arrival cost, Quadratic stage cost, Nonlinear systems.**

I. 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. Building on the success of moving horizon control, moving horizon estimation (*MHE*) has attracted the attention of researchers since the pioneering work of Jazwinski [1] (see also [2]–[4]). The interest in such estimation methods stems from the possibility of dealing with a limited amount of data, unlike the full information estimator (FIE), which takes into account all measurements available of the output of the system, increasing the size of the problem in every sampling time until it becomes computationally intractable. Besides, since MHE is optimisation based, it allows to incorporate constraints in a natural way.

MHE solves at each sampling time a finite horizon state estimation problem. When new measurements become available the old one is discarded from the estimation window. The information which are not included in the estimation window is summarised into the objective function through an extra term called AC. A good approximation of the AC allows to reduce the size of the estimation window and to have a performance comparable to that of the FIE. The most accepted way of approximating the AC is using a weighted $2-$norm of the states at the beginning of the estimation window [3], [5], [6], [7], [8], [9].

In recent years, both theoretical properties of several *MHE* schemes, including efficient computational methods, have been studied (see [9]–[14]).

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, [4] established the asymptotic stability of the estimation error for the least-square cost function. Furthermore, for convergent disturbances, the estimation error is also convergent [15]–[17]. 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 [11], [18]. The review made by Rawlings and Li [16] provides a general view of the problem relying on incremental input-output-to-state stability for detectability [19] 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. [20] identifies a broad class of cost functions that ensures *RGAS* of FIEs. The implication of *RGAS* on *MHE* was further investigated in [21] based on the results of [22]. Moreover, [21] 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 [21]. Finally, the necessary assumptions on the cost function were generalized in [23].

The results described for nonlinear detectable systems subject to bounded disturbances show a number of drawbacks. In particular, the disturbances gains obtained in [21] 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 [24], 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. In [25], the authors propose time-discounted schemes for FIE and MHE relying on the novel concept of time-discounted incrementally input-output-to-state-stable (i-ioss).

Another endeavour to obtain less conservative results is presented in [deniz2021multiplemodels], where several equivalence constants are introduced in order to avoid conservative bounds in both the estimation window length and the estimation error. Moreover, the equations probing robust stability for the case of quadratic SC in combination with adaptive AC stem for optimal design in the sense of minimising the estimation error.

II. Preliminaries and setup

A. Notation

Let $\mathbb{Z}$ denotes the integer numbers, $\mathbb{Z}_{[a,b]}$ denotes the set of integers in the interval $[a,b]$, with $b>a$ and $\mathbb{Z}_{\geq a}$ denotes the set of integers greater or equal to a. Boldface symbols denote sequences of finite or infinite length. We denote $\hat{x}_{j|k}$ as the state at time j estimated at time k. By $|x|$ we denote the euclidean norm of a vector $x \in \mathbb{R}^{n_x}$. Let $\|\boldsymbol{x}\| := \sup_{k\in\mathbb{Z}_{\geq 0}} |x_k|$ denote the supreme norm of the sequence $\boldsymbol{x}$ and $\|\boldsymbol{x}\|_{[a,b]} := \sup_{k\in\mathbb{Z}_{[a,b]}} |x_k|$, whereas x^T denotes transpose of vector x. 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 ζ is continuous, decreasing and $\lim_{t\to\infty} \zeta(t) = 0$. A function $\beta : \mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ 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 $r\in\mathbb{R}_{\geq 0}$. When necessary, we will use the notation $x_{i,k}^{(1)}$ and $x_{i,k}^{(2)}$ to differentiate i–th component of the states vector at time k of two discrete-time trajectories of the system, with $i\in\mathbb{Z}_{[1,n_x]}$. Moreover, $x_k^{(1)}(x_0^{(1)}, \boldsymbol{w}^{(1)})$ will denote a trajectory with initial condition $x_0^{(1)}$ and perturbed by the sequence $\boldsymbol{w}^{(1)}$. A similar notation is used for the case of continuous time systems, where t is used instead k to denote continuous time.

B. Problem statement

Let us consider a nonlinear discrete-time system with the following behaviour $\forall\, k \in \mathbb{Z}_{\geq 0}$

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

where $x_k \in \mathcal{X} \subset \mathbb{R}^{n_x}$ is the system state, $u_k \in \mathcal{U} \subset \mathbb{R}^{n_u}$ is the control action, assumed to be known and entering affinely to the system, $w_k \in \mathcal{W} \subset \mathbb{R}^{n_x}$ is the unmeasured additive process disturbance, $y_k \in \mathcal{Y} \subset \mathbb{R}^{n_y}$ is the output measurements and $v_k \in \mathcal{V} \subset \mathbb{R}^{n_v}$ is the measurement noise. The sets $\mathcal{X}, \mathcal{U}, \mathcal{W}, \mathcal{Y}$ and $\mathcal{V}$ are known to be compact and convex, with the null vector $\mathbf{0}$ of appropriate dimension in their interior. In the following we will assume that $f : \mathbb{R}^{n_x}\times\mathbb{R}^{n_u} \to \mathbb{R}^{n_x}$ and the output function $h : \mathbb{R}^{n_x} \to \mathbb{R}^{n_y}$ are both known Lipschitz functions in all their respective domains. Our goal is to compute an estimate $\hat{x}_{k|k}$ of $x_{k|k}$ as accurate as possible.

C. Moving horizon estimation scheme

In order to compute an estimate $\hat{x}_{k|k}$ of x_k, we solve the following moving horizon estimation problem at each samping time

$$
\begin{aligned}
\min_{\hat{x}_{k-N|k},\,\hat{\boldsymbol{w}}} \Psi := \Gamma_{k\,-\,N_e|k}(\chi) + \sum_{j=k-N_e}^{k-1} \ell_w\left(\hat{w}_{j|k}\right) + \\
\sum_{j=k-N_e}^{k} \ell_v\left(\hat{v}_{j|k}\right)
\end{aligned} \tag{2}
$$

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

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 both at time k based on measurements $\boldsymbol{y}_{[k-N_e,k]}$ available at time k. The length of the estimation window is denoted as N_e. The prior estimate $\hat{x}_{k-N_e|k}$ and the process noise $\hat{\boldsymbol{w}}_{[k-N_e|k]}$ are the optimization variables. The stage costs $\ell_w\left(\hat{w}_{j|k}\right)$ and $\ell_v\left(\hat{v}_{j|k}\right)$ penalises the estimated process disturbance $\hat{w}_{j|k}$ and the estimation residuals $\hat{v}_{j|k} = y_j - h\left(\hat{x}_{j|k}\right)$. Since in this work we deal with quadratic stage cost, $\ell_w\left(\hat{w}_{j|k}\right) = \hat{w}_{j|k}^T\, Q\, \hat{w}_{j|k}$ and $\ell_v\left(\hat{v}_{j|k}\right) = \hat{v}_{j|k}^T\, R\, \hat{v}_{j|k}$. The arrival cost $\Gamma_{k-N_e|k}(\chi)$ penalises the deviation of prior estimate $\hat{x}_{k-N_e|k}$ respect to the prior known $\bar{x}_{k-N_e}$ with the weighting matrix P^{-1}, i.e., $\Gamma_{k-N_e|k}(\chi) = \chi^T\, P^{-1}\, \chi$. In this work, the matrix P is given with the following behaviour [9]

$$
P_{k-N_e|k} = \begin{cases}
\frac{1}{\alpha_k} W_k & \text{if } \frac{1}{\alpha_k} \mathrm{Tr}\left(W_k\right) \leq c, \\
W_k & \text{otherwise},
\end{cases} \tag{3}
$$

with

$$
\begin{aligned}
\alpha_k &= 1 - \frac{1}{M_k}, \\
M_k &= \left[1 + \hat{x}_{k-N_e|k-1}^T\, P_{k-N_e-1|k-1} \hat{x}_{k-N_e|k-1}\right] \frac{\sigma}{\left|\hat{v}_{k-N_e|k}\right|^2}, \\
\hat{v}_{k-N_e|k} &= y_{k-N_e} - \hat{y}_{k-N_e|k}, \\
W_k &= P_{k-N_e-1|k-1} - \\
&\frac{P_{k-N_e-1|k-1}\hat{x}_{k-N_e|k-1}\hat{x}_{k-N_e|k-1}^T}{1 + \hat{x}_{k-N_e|k-1}^T P_{k-N_e-1|k-1}\hat{x}_{k-N_e|k-1} P_{k-N_e-1|k-1}},
\end{aligned}
$$

where $\sigma,\, c,\, \lambda \in R_{>0}$ are tuning parameters, $c>\lambda$, $P_0 = \lambda I_{n\times n}$ and $\sigma \gg \max\left\{\sigma_w^2,\, \sigma_v^2\right\}$, where σ_w^2 and σ_v^2 denote the process and measurement noise variances. The prior knowledge of the window $\bar{x}_{k-N}$ is updated using a smoothed estimate [26]

$$
\bar{x}_{k-N_e} = \hat{x}_{k-N_e|k-1}. \tag{4}
$$

III. Theoretical properties

The theoretical properties of an estimator with the behaviour exposed here, i.e., adaptive arrival cost given by (3) and quadratic stage costs were developed first in [27] and then generalised in [deniz2021multiplemodels] for the case of model uncertainties. The most remarkable result from those works, is that for a large enough window length, the estimator is robustly stable, whenever, the system being detectable, i.e., i-ioss. For the sake of completeness, the following definitions is given

Definition 1 system (1) is incrementally input-output-to-state stable (i-IOSS) if there exist functions $\beta \in \mathcal{KL}$ and $\gamma_w, \gamma_v \in \mathcal{K}$ such that for each pair of initial conditions $x_0^{(1)}, x_0^{(2)} \in \mathbb{R}^{n_x}$ and each two disturbance sequences $\boldsymbol{w}^{(1)}$ and $\boldsymbol{w}^{(2)}$ the following holds $\forall\, k \in \mathbb{Z}_{\geq 0}$

$$\begin{aligned} |x_k^{(1)} - x_k^{(2)}| \leq \quad & \beta\left(|x_0^{(1)} - x_0^{(2)}|, k\right) + \\ & \gamma_w\left(\|\boldsymbol{w}^{(1)} - \boldsymbol{w}^{(2)}\|_{[0,k-1]}\right) + \\ & \gamma_v\left(\|h\left(\boldsymbol{x}^{(1)}\right) - h\left(\boldsymbol{x}^{(2)}\right)\|_{[0,k-1]}\right). \end{aligned} \tag{5}$$

Then, if one implement a moving horizon estimator, it will be robustly stable whenever the length of the estimation window N_e being chose to satisfy $N_e \geq \mathcal{N}_e$, where $\mathcal{N}_e$ is shortest window for which the MHE is RGAS, an can be computed as follow

$$\begin{aligned} & \min \quad \mathcal{N}_e \\ & \text{s.t.} \left\{ \frac{\bar{e}_{x_0}^{\zeta}}{\mathcal{N}_e^{\iota}} k_{\beta_x}(\mathcal{N}_e) \leq \frac{\bar{e}_{x_0}}{\Pi_{N_e}}, \right. \end{aligned} \tag{6}$$

where, $\bar{x}_{x_0}$ is the maximum error on the initial condition of the system, $\Pi_{N_e} > 1$ is the decreasing rate of the transient due to start the estimation process from $\bar{x}_0 \neq x_0$, ζ and ι are constants which depend on the system and $k_{\beta_x}(\mathcal{N}_e)$ is a function evaluated at $\mathcal{N}_e$, and we will give further details later.

The estimation error bound for our moving horizon estimation is given by the following equation

$$|x_k - \hat{x}_{k|k}| \leq \quad \bar{\beta}_x\left(|x_0 - \bar{x}_0|, k\right) + \pi_w\left(\|\boldsymbol{w}\|\right) + \pi_v\left(\|\boldsymbol{v}\|\right), \tag{7}$$

where $\bar{\beta} \in \mathcal{KL}$, π_w, $\pi_v \in \mathcal{K}$ with the following behaviour

$$\begin{aligned} \bar{\beta}_x\left(|x_0 - \bar{x}_0|, k\right) \leq \quad & \left(\frac{\Pi_{N_e} + \mu}{\Pi_{N_e}(1+\mu)}\right)^i \frac{|x_0 - \bar{x}_0|^{\zeta}}{j^{\iota}} k_{\beta_x}(j), \\ \pi_w\left(\|\boldsymbol{w}\|\right) \leq \quad & \frac{\Pi_{N_e}(1+\mu)}{\Pi_{N_e} - 1} \|\boldsymbol{w}\|^{\zeta} k_w(j), \\ \pi_v\left(\|\boldsymbol{v}\|\right) \leq \quad & \frac{\Pi_{N_e}(1+\mu)}{\Pi_{N_e} - 1} \|\boldsymbol{v}\|^{\zeta} k_v(j), \end{aligned} \tag{8}$$

with $k = i(N_e + 1) + j$, where $i = \lfloor k/(N_e + 1) \rfloor$, $j = \bmod\ (k, N_e + 1)$, $\Pi_{N_e} \in \mathbb{R}_{>1}$ is the decreasing rate of the error on the initial condition, which is proportional to the length of the horizon window N_e and $\mu \in \mathbb{R}_{>0}$ is a parameter which determines the volume of the invariant space for the estimation error. The functions $k_{\beta_x}(j)$, $k_w(j)$ and $k_v(j)$ are given by the following expressions

$$\begin{aligned} k_{\beta_x}(j) := \quad & \frac{e_\beta c_\beta}{j^{q-\iota}} \left(1 + \frac{p e_\beta \bar{\lambda}_P^{p_\beta/2}}{\underline{\lambda}_P^{p_\beta}}\right) + \frac{e_{\gamma_3} c_{\gamma_3} p_{\gamma_3} \bar{\lambda}_P^{\gamma_3/2}}{j^{\gamma_3/2-\iota} \underline{\lambda}_Q^{\gamma_3}} + \\ & \frac{e_{\gamma_4} c_{\gamma_4} p_{\gamma_4} \bar{\lambda}_P^{\gamma_4/2}}{j^{\gamma_4/2-\iota} \underline{\lambda}_R^{\gamma_4}}, \\ k_w(j) := \quad & \frac{e_\beta c_\beta\, p e_\beta \bar{\lambda}_Q^{p_\beta/2}}{j^{q-p_\beta/2} \underline{\lambda}_P^{p_\beta}} + e_{\gamma_3} c_{\gamma_3} + \frac{e_{\gamma_3} c_{\gamma_3} p_{\gamma_3} \bar{\lambda}_Q^{\gamma_3/2}}{\underline{\lambda}_Q^{\gamma_3}} + \\ & \frac{e_{\gamma_4} c_{\gamma_4} p_{\gamma_4} \bar{\lambda}_Q^{\gamma_4/2}}{\underline{\lambda}_R^{\gamma_4}}, \\ k_v(j) := \quad & \frac{e_\beta c_\beta\, p e_\beta \bar{\lambda}_R^{p_\beta/2}}{j^{q-p_\beta/2} \underline{\lambda}_P^{p_\beta}} + e_{\gamma_4} c_{\gamma_4} + \frac{e_{\gamma_3} c_{\gamma_3} p_{\gamma_3} \bar{\lambda}_R^{\gamma_3/2}}{\underline{\lambda}_Q^{\gamma_3}} + \\ & \frac{e_{\gamma_4} c_{\gamma_4} p_{\gamma_4} \bar{\lambda}_R^{\gamma_4/2}}{\underline{\lambda}_R^{\gamma_4}}, \end{aligned} \tag{9}$$

with $k_{\beta_x}(0) = \delta k_{\beta_x}(1)$, $k_w(0) = \delta k_w(1)$ and $k_v(0) = \delta k_v(1)$, for some $\delta > 1$. The quantities involved in (9) depends on constants of the i-ioss function (which depend at the same time of the dynamic of the system) and the eigenvalues of the arrival and stage cost matrices. Based on (9), one could compute the optimal values of the maximum and minimum eigenvalues of each matrix, i.e., $\bar{\lambda}_P$, $\underline{\lambda}_P$, $\bar{\lambda}_Q$, $\underline{\lambda}_Q$, $\bar{\lambda}_R$ and $\underline{\lambda}_R$. However, since matrix P is updated at every sampling time, $\bar{\lambda}_P$, $\underline{\lambda}_P$ are updated, as well. Moreover, from (3) one can see that for the case of vanishing disturbances, i.e., $\lim_{k\to\infty} v_k = 0$ and $\lim_{k\to\infty} w_k = 0$, $\lim_{k\to\infty} |P_k - P_{k-1}| = 0$, i.e., matrix P stop being updated. However, this is not the case for the case of non vanishing disturbances (a more realistic one) and the eigenvalues $\bar{\lambda}_P$ and $\underline{\lambda}_P$ will change for all times, specially when constant c in (3) is not properly chose. In order to overcome this problem and at the same time improving the estimation process based on previous results regarding robust stability, we introduce a modification to (3) which consist of scaling matrix P in order to preserve the conditioning number, allowing at the same time a change in the values of $\bar{\lambda}_P$ and $\underline{\lambda}_P$.

IV. Proposed scheme

A. Computing optimal eigenvalues

The set of equations in (9) can be slightly simplified to formulate the following optimisation problem and compute the optimal eigenvalues

$$\begin{aligned} & \min_{\underline{\lambda}_P, \bar{\lambda}_P, \underline{\lambda}_Q, \bar{\lambda}_Q, \underline{\lambda}_R, \bar{\lambda}_R} \quad \mathcal{J} \\ & \text{s.t.} \left\{ \begin{array}{l} 0 < \underline{\lambda}_P \leq \bar{\lambda}_P, \\ 0 < \underline{\lambda}_Q \leq \bar{\lambda}_Q, \\ 0 < \underline{\lambda}_R \leq \bar{\lambda}_R, \end{array} \right. \end{aligned} \tag{10}$$

where the objective function is defined as

$$\mathcal{J} := \quad k_{\beta_x}(0) + k_w(0) + k_v(0). \tag{11}$$

Once the eigenvalues were determined, the matrices P^{-1}, Q and R can be designed with the computed conditioning number.

B. Scaling the arrival cost

The matrices of the stage cost are initialised with the values computed former, whereas the matrix of the arrival cost can be initialised with the values computed or as in [9]. In every sampling time, the arrival cost is updated according (3), then the matrix P is scaled to preserve the conditioning number computed with (10). In order to scale P, a slightly modified version of the optimization problem published in [28], page 37, is solved at each sampling-time

$$\begin{aligned} &\text{find} \quad M,\, N \\ &\text{s.t.} \begin{cases} \underline{\lambda}_P^2\, M \le P^{-T}\, N\, P^{-1} \le C_P^2\, \underline{\lambda}_P^2\, M, \\ M > 0, \\ N > 0, \end{cases} \end{aligned} \tag{12}$$

where $M,\, N \in \mathbb{R}^{n_x \times n_x}$ and $C_P = \bar{\lambda}_P / \underline{\lambda}_P$ is the conditioning number. Note that (12) does not find an optimal solution, just a feasible one, since the condition number is given, in comparison with [28] where the conditioning number is minimised. After solving (12), the matrix P is scaled as $N_{chol}\, P\, M_{chol}^{-1}$, where N_{chol} and M_{chol} are the Cholesky factorisation of matrices N and M, respectively. Another approach consist of solving the following alternative problem

$$\begin{aligned} &\text{find} \quad a_i \\ &\text{s.t.} \left\{ \lambda_j^{n_x} + a_{n_x-1}\lambda_j^{n_x-1} + \ldots + a_0 = 0, \right. \end{aligned} \tag{13}$$

with $\underline{\lambda} = 1/\bar{\lambda}_P$ and $\bar{\lambda} = \underline{\lambda} C_P$, whereas the eigenvalues between $\underline{\lambda}$ and $\bar{\lambda}$ should be properly chosen or scaled from those ones after compute (3). Then, matrix P is reconstructed from constants a_i.

V. Simulations and results

A. Example 1

In order to evaluate and compare the performance of the proposed MHE, let us consider as a first example the following linear discrete-time system

$$\begin{aligned} x_{k+1} = &\begin{bmatrix} 0.8\, x_{1,k} + 0.2\, x_{2,k} + 0.5\, w_{1,k} \\ -0.3\, x_{1,k} + 0.5\, x_{2,k} + w_{2,k} \end{bmatrix}, \\ y_k = &\; x_{2,k} + v_k \end{aligned} \tag{14}$$

The system is affected by process noise with uniform distribution $w_{1,k} \in [-0.005,\, 0.005]$ and $w_{2,k} \in [-0.05,\, 0.05]$ and measurement noise with normal distribution $v_k \sim \mathcal{N}\left(0,\, 0.01^2\right)$. The state, process and residual estimated are constrained to the following set

$$\begin{aligned} \mathcal{X} &:= \{x : |x_1| \le 10,\, |x_2| \le 10\}, \\ \mathcal{W} &:= \{w : |w_1| \le 0.03,\, |w_2| \le 0.3\}, \\ \mathcal{V} &:= \{v : |v| \le 0.03\}, \end{aligned}$$

The system is i-ioss with the following function

$$|x_k^{(1)} - x_k^{(2)}| \le \quad \tfrac{2\times 22.86^{2.037}}{k^{1.0186}} + 0.25\, \|\boldsymbol{w}\| + 0.25\, \|\boldsymbol{v}\|. \tag{15}$$

whenever $x \in \mathcal{X}$ and robustly stable for $N_e \ge 9$ with the following equation bounding the estimation error

$$|x_k - \hat{x}_{k|k}| \le \quad \tfrac{0.976^i \times 4.1}{\sqrt{j}} + 1.07\, \|\boldsymbol{w}\|^{2.037} + 0.59\, \|\boldsymbol{v}\|^{2.037} \tag{16}$$

Three MHE with $N_e = 9$ will be implemented to compare their performance. Besides, two set o matrices will be implemented for each estimator. The first estimator is a standard NMHE ($NMHE_{adap}$) with quadratic stage cost and adaptive arrival cost updated according (3). The first set of matrices correspond to those of optimal values computed from (10).

Set 1

$$P_{0_1} = \begin{bmatrix} 10^6 & 0 \\ 0 & 10^6 \end{bmatrix}, Q_1 = \begin{bmatrix} 900194 & 0 \\ 0 & 617831 \end{bmatrix}, R_1 = 549935,$$

Set 2

$$P_{0_1} = \begin{bmatrix} 10^6 & 0 \\ 0 & 10^6 \end{bmatrix}, Q_1 = \begin{bmatrix} 100 & 0 \\ 0 & 10 \end{bmatrix}, R_1 = 100,$$

The second ($NMHE_{fix\,opt}$) is a NMHE with fixed P, Q and R matrices with optimal eigenvalues computed offline according (10). For this estimator, only 1 set is implemented

$$P_{0_2} = \begin{bmatrix} 7.77 \times 10^{-7} & 0 \\ 0 & 1.37 \times 10^{-6} \end{bmatrix},$$

$$Q_2 = \begin{bmatrix} 900194 & 0 \\ 0 & 617831 \end{bmatrix}, R_2 = 549935,$$

The third estimator (MHE_{scaled}) is a NMHE similar to the first, but the matrix P is updated according (3) and then to (13). Both sets of matrices are the same of those of $NMHE_{adap}$. The performance of the estimators $NMHE_{fix\,opt}$ and $NMHE_{scaled}$ will be evaluated according the metric $100 \times (1 - err_{estimator}/err_{adap})$.

TABLE I

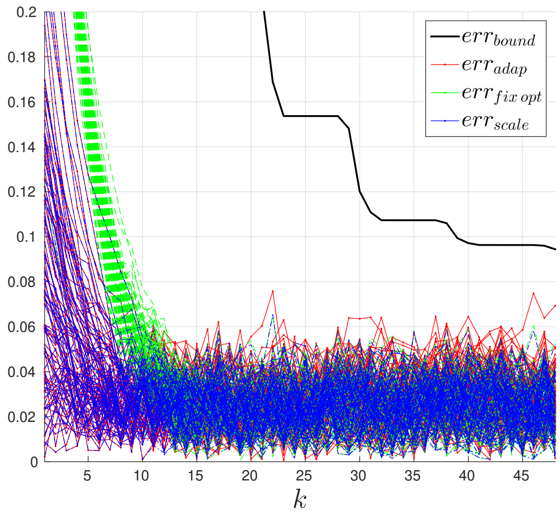
	$NMHE_{fix\,opt}$	$NMHE_{scaled}$
set1	$-88.30\,\%$	$7.06\,\%$
set2	$-88.30\,\%$	$22.21\,\%$

Table I summarises the averaged performance about 100 trials. The performance of $NMHE_{fix\,opt}$, in which all matrices are initialised with the values computed from (10) and does not update the arrival cost, is deteriorated in comparison with the reference estimator $NMHE_{adap}$, whereas the estimator $NMHE_{scaled}$ which scales the arrival cost matrix after updating it according (3) improves its performance with both set of matrices.

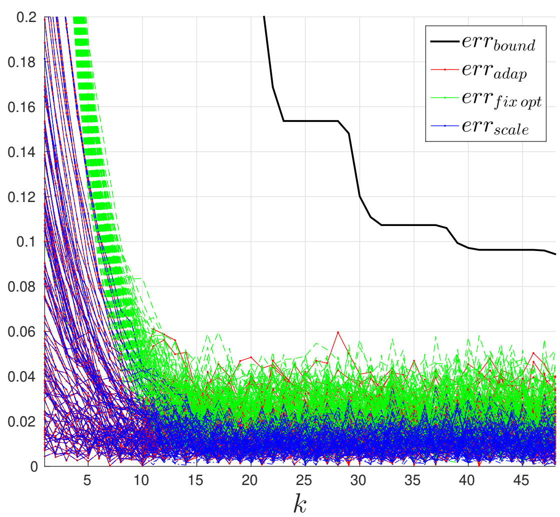
B. Example 2

As a second example, let us consider the discrete-time nonlinear system

$$\begin{aligned} x_{k+1} = &\begin{bmatrix} 0.8\, x_{1,k} + 0.2\, x_{2,k}^3 + 0.5\, w_{1,k} \\ \frac{-0.3\, x_{1,k}}{1+x_{2,k}} + 0.5\cos(x_{2,k}^2) + w_{2,k} \end{bmatrix}, \\ y_k = &\; x_{2,k} + v_k \end{aligned} \tag{17}$$



(a)



(b)

Fig. 1: Estimation error bound (black) and errors committed by MHE_{adap} (red), $MHE_{fix\,opt}$ (green) and MHE_{scaled} (blue) for a) matrices from set 1. b) Matrices from set 2.

The constraints of the estimators are the same is Example 1, as well as the noises. The systems is i-ioss with the following function

$$|x_k^{(1)} - x_k^{(2)}| \leq \frac{2\times 16.28^{2.448}}{k^{1.224}} + 0.25\,\|\boldsymbol{w}\| + 0.25\,\|\boldsymbol{v}\|. \tag{18}$$

and robustly stable for $N_e \geq 9$, and the function which bounds the estimation error is the following

$$|x_k - \hat{x}_{k|k}| \leq \frac{0.976^{i}\times 5.46}{\sqrt{j}} + 1.08\,\|\boldsymbol{w}\|^{2.448} + 0.554\,\|\boldsymbol{v}\|^{2.448} \tag{19}$$

The set of matrices for the estimator $NMHE_{adap}$ are the followings

Set 1

$$P_{0_1} = \begin{bmatrix} 10^6 & 0 \\ 0 & 10^6 \end{bmatrix}, Q_1 = \begin{bmatrix} 1.46\times 10^6 & 0 \\ 0 & 1.09\times 10^6 \end{bmatrix},$$
$$R_1 = 492541,$$

Set 2

$$P_{0_1} = \begin{bmatrix} 10^6 & 0 \\ 0 & 10^6 \end{bmatrix}, Q_1 = \begin{bmatrix} 100 & 0 \\ 0 & 10 \end{bmatrix}, R_1 = 100,$$

Set 1 and 2 are the same for the $NMHE_{fix\,opt}$

$$P_{0_2} = \begin{bmatrix} 1.1\times 10^{-6} & 0 \\ 0 & 6.66\times 10^{-6} \end{bmatrix},$$
$$Q_2 = \begin{bmatrix} 1.46\times 10^6 & 0 \\ 0 & 1.09\times 10^6 \end{bmatrix}, R_2 = 492541,$$

and the sets for the estimator $NMHE_{scaled}$ are the same as those for $NMHE_{adap}$.

TABLE II

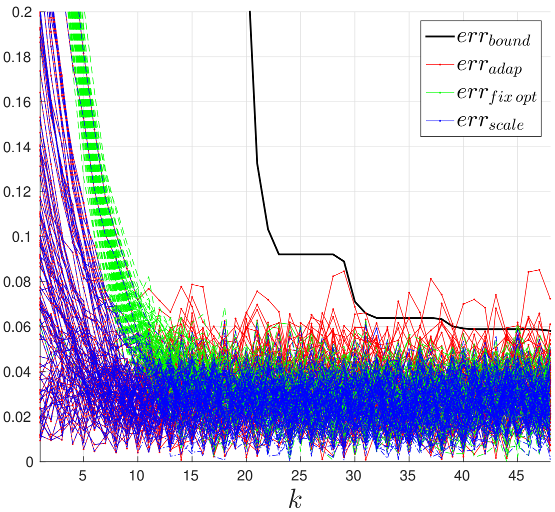
	$NMHE_{fix\,opt}$	$NMHE_{scaled}$
set1	$-76.00\,\%$	$12.71\,\%$
set2	$-76.00\,\%$	$25.45\,\%$

VI. Conclusion

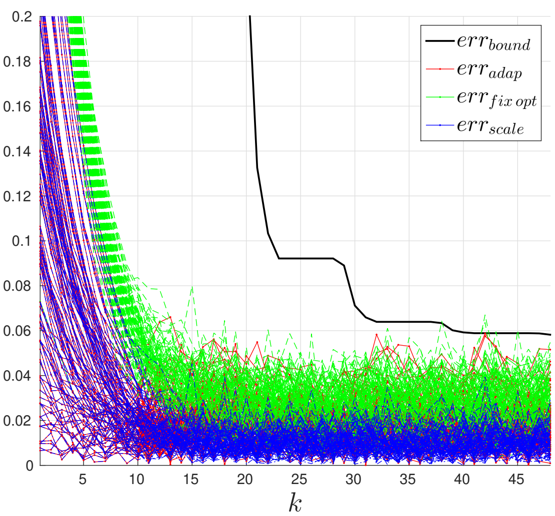
In this paper, we have shown how can improve the estimation of a moving horizon estimator by controlling the conditioning number of the arrival cost weighting matrix. The key idea supporting the results comes from previous results on the robust stability of an MHE. The equations which guarantee robust stability for an MHE with adaptive arrival cost and quadratic stage cost shows a clear relationship between the eigenvalues of the weighting matrices of the MHE. Maximum and minimum eigenvalues were treated as optimisation variables in order to minimise the error bound. We have tested two approaches. The first consist of computing the optimal eigenvalues and design the weighting matrix from those values, remaining them unchanged during the estimation. In the second approach, the conditioning number of the arrival cost weighting matrix computed from the optimal eigenvalues was controlled in every sampling time in order to maintain its value, allowing a change in the eigenvalues itself, while the stage cost matrices were given other values rather than those computed as optimal. The latter method resulted in higher performance in the estimations of states of linear and nonlinear systems. However, both methods are shown to behave according to the equations which govern the robust stability. As expected, both simulation studies put on evidence the relevance of updating the arrival cost of an estimator with a moving horizon scheme.

Acknowledgment

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



(a)



(b)

Fig. 2: Estimation error bound (black) and errors committed by MHE_{adap} (red), $MHE_{fix\,opt}$ (green) and MHE_{scaled} (blue) for a) matrices from set 1. b) Matrices from set 2.

References

[1] A. Jazwinski, "Limited memory optimal filtering," *IEEE Transactions on Automatic Control*, vol. 13, no. 5, pp. 558–563, 1968.

[2] F. C. Schweppe, *Uncertain dynamic systems*. Prentice Hall, 1973.

[3] C. V. Rao, J. B. Rawlings, and J. H. Lee, "Constrained linear state estimation—a moving horizon approach," *Automatica*, vol. 37, no. 10, pp. 1619–1628, 2001.

[4] C. V. Rao, J. B. Rawlings, and D. Q. Mayne, "Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations," *IEEE transactions on automatic control*, vol. 48, no. 2, pp. 246–258, 2003.

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

[6] P. Kühl, M. Diehl, T. Kraus, J. P. Schlöder, and H. G. Bock, "A real-time algorithm for moving horizon state and parameter estimation," *Computers & chemical engineering*, vol. 35, no. 1, pp. 71–83, 2011.

[7] E. Chu, A. Keshavarz, D. Gorinevsky, and S. Boyd, "Moving horizon estimation for staged qp problems," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE, 2012, pp. 3177–3182.

[8] K. Baumgärtner, A. Zanelli, and M. Diehl, "A gradient condition for the arrival cost in moving horizon estimation," in *2020 European Control Conference (ECC)*. IEEE, 2020, pp. 1286–1291.

[9] G. Sánchez, M. Murillo, and L. Giovanini, "Adaptive arrival cost update for improving moving horizon estimation performance," *ISA transactions*, vol. 68, pp. 54–62, 2017.

[10] A. Alessandri, M. Baglietto, and G. Battistelli, "Robust receding-horizon state estimation for uncertain discrete-time linear systems," *Systems & Control Letters*, vol. 54, no. 7, pp. 627–643, 2005.

[11] ——, "Moving-horizon state estimation for nonlinear discrete-time systems: New stability results and approximation schemes," *Automatica*, vol. 44, no. 7, pp. 1753–1765, 2008.

[12] ——, "Min-max moving-horizon estimation for uncertain discrete-time linear systems," *SIAM Journal on Control and Optimization*, vol. 50, no. 3, pp. 1439–1465, 2012.

[13] 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*, vol. 13, no. 6, pp. 653–664, 2016.

[14] H. Sartipizadeh and T. L. Vincent, "Computationally tractable robust moving horizon estimation using an approximate convex hull," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE, 2016, pp. 3757–3762.

[15] J. B. Rawlings and D. Q. Mayne, *Model predictive control: Theory and design*. Nob Hill Pub. Madison, Wisconsin, 2009.

[16] J. B. Rawlings and L. Ji, "Optimization-based state estimation: Current status and some new results," *Journal of Process Control*, vol. 22, no. 8, pp. 1439–1444, 2012.

[17] D. A. Allan and J. B. Rawlings, "Moving horizon estimation," in *Handbook of Model Predictive Control*. Springer, 2019, pp. 99–124.

[18] A. Alessandri, M. Baglietto, G. Battistelli, and V. Zavala, "Advances in moving horizon estimation for nonlinear systems," in *Decision and Control (CDC), 2010 49th IEEE Conference on*. IEEE, 2010, pp. 5681–5688.

[19] E. D. Sontag and Y. Wang, "Output-to-state stability and detectability of nonlinear systems," *Systems & Control Letters*, vol. 29, no. 5, pp. 279–290, 1997.

[20] W. Hu, L. Xie, and K. You, "Optimization-based state estimation under bounded disturbances," in *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE, 2015, pp. 6597–6602.

[21] M. A. Müller, "Nonlinear moving horizon estimation in the presence of bounded disturbances," *Automatica*, vol. 79, pp. 306–314, 2017.

[22] L. Ji, J. B. Rawlings, W. Hu, A. Wynn, and M. Diehl, "Robust stability of moving horizon estimation under bounded disturbances," *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3509–3514, 2016.

[23] W. Hu, "Robust stability of optimization-based state estimation," *arXiv preprint arXiv:1702.01903*, 2017.

[24] S. Knüfer and M. A. Müller, "Robust global exponential stability for moving horizon estimation," in *2018 IEEE Conference on Decision and Control (CDC)*, 2018, pp. 3477–3482.

[25] S. Knuefer and M. A. Mueller, "Nonlinear full information and moving horizon estimation: Robust global asymptotic stability," *arXiv preprint arXiv:2105.02764*, 2021.

[26] P. K. Findeisen, "Moving horizon state estimation of discrete time systems," Ph.D. dissertation, University of Wisconsin–Madison, 1997.

[27] N. Deniz, M. Murillo, G. Sanchez, and 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.

[28] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. SIAM, 1994.