Moving Horizon Estimator with filtering and adaptive sampling

Federico Oliva * **Daniele Carnevale** *

* *Dipartimento di Ing. Civile e Ing. Informatica, University of Rome "Tor Vergata"*

Abstract: Optimisation based algorithms known as *Moving Horizon Estimator* (MHE) have been developed through the years. In this work, we propose two solutions to decrease the computational cost of MHE, limiting its applicability in real-time applications. The proposed solutions rely on output filtering and adaptive sampling. The use of filters reduces the total amount of data by shortening the length of the moving window (buffer) and consequently decreasing the time consumption for plant dynamics integration. The proposed adaptive sampling policy allows for discarding data that do not yield significant improvements in the estimation error. Simulations on several cases are provided to corroborate the effectiveness of the proposed strategies.

Keywords: Adaptive observer, Nonlinear systems, Moving Horizon Estimator

1. INTRODUCTION

A well-known problem in estimation theory is the design of robust observers, intending to reconstruct the plant states that are not directly measurable by sensors. Observer design has been widely studied and developed for linear and nonlinear systems. The *Luenberger observer* and *Kalman filter* (Davis, 2002; Thrun et al., 2006) are classic solutions to the observation problem on linear systems. There are extensions to systems described by nonlinear dynamics, such as the *Extended Kalman filter* (Thrun et al., 2006; Reif et al., 1999) or the *Unscented Kalman filter* (Thrun et al., 2006) based on invariant manifolds (Karagiannis and Astolfi, 2005; Karagiannis et al., 2008). In particular, the approaches above are also used to estimate plant parameters (adaptive observers). In this field of research, many results have been provided by scientists, and solutions have been tailored for specific classes of nonlinear systems (Luders and Narendra, 1974; Krener and Isidori, 1983; Marino, 1990; Tyukin et al., 2013; Marine et al., 2001; Marino and Tomei, 1992). In this work, we deal with the class of *Moving Horizon Estimators* (MHE) (Michalska and Mayne, 1995; Kang, 2006; Wynn et al., 2014; Sui et al., 2010; Suwantong et al., 2014; Schiller et al., 2021). In MHE, the estimation problem is re-cast into an optimisation problem. In particular, in (P.E. Moraal, 1995) the Newton method is exploited to minimise the mismatch among measurements and estimated output using results from (Aeyels, 1981) and (Broyden, 2000). Observability and persistent excitation conditions for MHE observers with nonlinear plants are recalled in (Glad, 1983). Due to its optimisation-based approach, the main drawback of the MHE is the computational cost. Then, the goal of this work is to propose strategies to decrease the computational time required to run MHE observers. The first proposed solution to tackle the issue hinges upon output filtering. A general analysis in terms of observability of such an approach can be found in (Menini et al., 2019, 2022). Furthermore, in this work, an adaptive sampling policy is proposed to autonomously select the *"best"* output samples to be processed to reduce the state or parameter estimation error. The paper unfolds as follows: in Sec. 2 a general analysis of the MHE is provided, focusing on the *Standard* MHE and convergence conditions are recalled. In Sec.3, the proposed solutions to decrease computational costs are introduced. Lastly, simulation results are provided in Sec.4, and the computational-time speed-up factor is discussed. Conclusions and future work are provided in Sec. 5.

2. MOVING HORIZON ESTIMATORS

Consider a general nonlinear system in the following form:

$$\dot{\boldsymbol{\xi}} = f(\boldsymbol{\xi}, \boldsymbol{u}), \tag{1a}$$
$$\boldsymbol{y} = h(\boldsymbol{\xi}, \boldsymbol{u}), \tag{1b}$$

where $\boldsymbol{\xi} \in \mathbb{R}^n$ is the state vector, $\boldsymbol{u} \in \mathbb{R}^m$ is the control input, and $\boldsymbol{y} \in \mathbb{R}^p$ is the measured output. Following the notation introduced in (P.E. Moraal, 1995), we recall next the definition of moving horizon estimators (MHE).

2.1 MHE definition

In the rest of the paper we assume that inputs and outputs of plant (1) are sampled every T_s seconds, i.e. T_s is the sampling time of (1). We also consider a down-sampling of the I/O signals to be saved into the MHE buffer (moving window data), such that samples at time instants t_k are buffered, with $k \in \mathbb{Z}$, and $t_k - t_{k-1} = N_{T_s} \cdot T_s$ (multiple of sampling time) for some $N_{T_s} \in \mathbb{N}_{\geq 1}$. It is also assumed the input $\boldsymbol{u}$ to be constant within sampling time intervals, i.e. it is piece-wise constant. We now recall the concept of *N-lifted* system as defined in (Tousain et al., 2001), namely the procedure through which a continuous system is described by a discrete one with a finite number of samples (data buffer). Define the (down-sampled) output $\mathrm{Y}_k = [\boldsymbol{y}_{k-N+1}, \dots, \boldsymbol{y}_k]^T \in \mathbb{R}^{(N\times p)}$ and input $\mathrm{U}_k = [\boldsymbol{u}_{k-N+1}, \dots, \boldsymbol{u}_k]^T \in \mathbb{R}^{(N\times m)}$ vectors of (1), where $\boldsymbol{y}_j \triangleq$

$\boldsymbol{y}(t_j)$ and $\boldsymbol{u}_j = \boldsymbol{u}(t_j)$, and with $j \in \{k-N+1, .., k\}$. N is the number of samples defining the moving window data (buffer). The *N-lifted* system operator is defined as

$$\mathrm{H}_k(\boldsymbol{\xi}_{k-N+1}, \boldsymbol{u}(\cdot)) = \begin{bmatrix} h(\boldsymbol{\xi}_{k-N+1}, \boldsymbol{u}_{k-N+1}) \\ h(\phi(t_{k-N+2}, t_{k-N+1}, \boldsymbol{\xi}_{k-N+1}, \boldsymbol{u}(\cdot)), \boldsymbol{u}_{k-N+2}) \\ \vdots \\ h(\phi(t_k, t_{k-1}, \boldsymbol{\xi}_{k-1}, \boldsymbol{u}(\cdot)), \boldsymbol{u}_k) \end{bmatrix} \tag{2}$$

where $\phi(t_i, t_j, \boldsymbol{\xi}_j, \boldsymbol{u}(\cdot)) = \boldsymbol{\xi}(t_i) = \boldsymbol{\xi}_i$ is the state transition map of (1) from t_j to t_i, with $k-N+1 \leq j < i \leq k$ and the piece-wise continuous input $\boldsymbol{u}(\tau)$, $\tau \in [t_j, t_i]$, whose discontinuities corresponds to sampling times $t_j + qT_s$ with $q \in \mathbb{N}$ such that $0 < q < N_{T_s}(i-j)$ (input sampled with zero order hold devise with sampling time T_s). Within the framework of MHEs, the estimation problem consists in finding the solution, at each time step k, to the following minimisation problem:

$$\min_{\hat{\boldsymbol{\xi}}_{k-N+1}} \ \mathrm{V}_k(\mathrm{Y}_k, \mathrm{H}_k(\hat{\boldsymbol{\xi}}_{k-N+1}, \boldsymbol{u}(\cdot))), \tag{3}$$

for which we selected a quadratic function V_k defined as

$$V_k \triangleq \sum_{j=1}^{N} (\mathrm{Y}_k^j - \hat{\mathrm{H}}_k^j)^T W_j (\mathrm{Y}_k^j - \hat{\mathrm{H}}_k^j), \tag{4}$$

where we used the compact notation $\hat{H}_k = \mathrm{H}_k(\hat{\boldsymbol{\xi}}_{k-N+1}, \boldsymbol{u}(\cdot))$, $W_i \in \mathbb{R}^{p \times p}$ are symmetric and positive definite weight matrices, Y_k^j and $\hat{\mathrm{H}}_k^j$ are the *j-th* rows of the matrices Y_k and $\hat{H}_k$, respectively. Then, we define the MHE associated to the *N-lifted* system of (1) as the observer described by the following nonlinear dynamical system

$$\boldsymbol{\zeta}_{i+1} = \boldsymbol{\Psi}(\mathrm{Y}_k, \mathrm{H}_k(\boldsymbol{\zeta}_i, \boldsymbol{u}(\cdot)), \boldsymbol{\zeta}_i), i = 0 \dots K, \ K \in \mathbb{N}, \tag{5a}$$

$$\boldsymbol{\zeta}_0 = \hat{\boldsymbol{\xi}}_{k-N+1}, \tag{5b}$$

$$\hat{\boldsymbol{\xi}}_k = \phi(t_k, t_{k-N+1}, \boldsymbol{\zeta}_K, \boldsymbol{u}(\cdot)), \tag{5c}$$

i.e., the observer consists of the discrete time algorithm (5a) that is evaluated K times and aims at the minimisation of V_k between (sub)samples, for a given data window. As an example, $\boldsymbol{\Psi}$ could be the discrete time jump map yield by a gradient(-like) minimisation algorithm of (4). The new estimate $\boldsymbol{\zeta}_K$ of $\boldsymbol{\xi}_{k-N+1}$ is then propagated to obtain the estimate of $\boldsymbol{\xi}_k$ through (5c). The structure of this MHE outlines the one described in (Wynn et al., 2014), with the exception that we consider a *single-shooting* rather than a *simultaneous* MHE observer, i.e. the optimisation is performed on the variable $\hat{\boldsymbol{\xi}}_{k-N+1}$ only, and not simultaneously on the vector $[\hat{\boldsymbol{\xi}}_{k-N+1}, \dots, \hat{\boldsymbol{\xi}}_k]$. This yields the loss of constraints $||\hat{\boldsymbol{\xi}}_{j+1} - \phi(t_{j+1}, t_j, \hat{\boldsymbol{\xi}}_j, \boldsymbol{u}(\cdot))|| = 0$, $\forall j = \{k-N+1, \dots, k\}$, and leaves us with an unconstrained optimisation problem. Please note that we did not consider a *terminal cost* like in (Sui et al., 2010; Suwantong et al., 2014; Kühl et al., 2011), since forward propagation of the previous estimate is obtained by mean of (5c) as in (Wynn et al., 2014). To have (practical) convergence of the estimation error, we assume that (1) is *N-observable* according to the Definition provided in (Aeyels, 1981). This condition can be related to the *(N+1)-step* observability provided in (Wynn et al., 2014). Note that we are not considering measurement noise as part of the optimisation problem as in (Schiller et al., 2021). Therefore, in case of no disturbances and measurement noise, if K is sufficiently high, the MHE (5) provides an asymptotic estimate of $\boldsymbol{\xi}$. It is worth noticing that the solution of problem (3) coincides with solving the following system of equations:

$$\left\| \mathrm{Y}_k^j - \hat{\mathrm{H}}_k^j \right\| = 0, j = 1, \dots, N. \tag{6}$$

Without loss of generality, from now on, we will consider [1] $p = 1$. Therefore, (6) consists of exactly N conditions.

2.2 MHE implementation

MHE real-time implementation rises a major concern due to high computational cost. A common practice in implementing MHE consists in limiting the number of optimisation cycles, i.e. select sufficiently low K in (5) yielding estimation performance deterioration (Schiller et al., 2021), and letting convergence issues for the observer arise. Indeed, if (1) is linear and $\boldsymbol{\Psi}$ is the implementation of the Newton's method, since V_k is quadratic with respect to the estimation error, then the observer (5) provides exact estimation with just $K = 1$ becoming a *dead-beat* observer (under perfect measurements). However, in the general case, the higher K, the lower the estimation error, i.e. $e \triangleq \xi_{k-N+1} - \zeta_{k-N+1}$. Standard approaches to reduce the computational cost of MHEs balance convergence performances and bounds on the error dynamics. In particular, (Kang, 2006) proves the error dynamics to be bounded under technical assumptions on the smoothness of $(f(\cdot), h(\cdot))$, assuming that the plant is *(N+1)-step* observable. Hinging upon such results, we focus on the possibility of reducing the window (buffer) length N used to evaluate the cost function V_k. In fact, the higher N, the more model integrations have to be computed to evaluate $\hat{H}_k$ in (3). When the MHE is not invoked (i.e. out of the down-sampling instants t_i), the last estimated state is simply propagated ahead of T_s seconds numerically integrating (1). For a more detailed description of the algorithm implementation, refer to (Oliva and Carnevale, 2022). Among several convergence analysis results provided for MHE, we consider two sufficient conditions for selecting N. Recalling (Aeyels, 1981; P.E. Moraal, 1995; Kang, 2006), it holds:

- C1 : $N \geq 2n+1$ is a *generic* sufficient condition to find a solution to (3) with an optimisation algorithm (Newton-like in P.E. Moraal (1995)).
- C2 : $N \geq n$ is a necessary condition exploiting a one-step ($K = 1$) BFGS optimisation algorithm yielding local practical stability of the estimation error (Kang, 2006).

The necessary condition $N \geq n$ can also be retrieved in the standard (local) observability condition of (non)linear systems and is related to the the N constraints in (6). This set of equations needs to be considered to solve, at least locally, the minimisation problem (3) with unique solution. In the next, since the sufficient condition C2 hinges upon

[1] All the results obtained in this work can be generalised to the case in which $p > 1$, under some observability conditions. However, due to space restrictions, this analysis has not been reported in the present work.

technical assumptions reducing the family of systems for which condition C1 holds, we propose the use of output filtering and adaptive sampling to relax condition C1 and further reduce the length of the buffer N, i.e. decreasing the computational costs.

3. STRATEGIES FOR COMPUTATIONAL EFFICIENCY

In the next paragraphs, in order to motivate our approaches to decrease the computational time of the MHE (5), we highlight the role of the buffer length N and the down-sampling parameter N_{T_s} on two different models, namely the Van der Pol oscillator and a nonlinear system describing the dynamics of plasma waves in the presence of runaway electrons (Buratti et al., 2021; Carnevale et al., 2021).

3.1 Buffer length N

The goal of the forthcoming analysis is to show how the buffer length N influences the performance of the observer on a classical nonlinear system, namely the Van der Pol oscillator described by the following equations:

$$\dot{\xi}_1 = \xi_2, \tag{7a}$$
$$\dot{\xi}_2 = -\xi_1 + \eta(1-\xi_1^2)\xi_2, \tag{7b}$$
$$y = \xi_1, \tag{7c}$$

where $y \in \mathbb{R}$ is the scalar output, $\boldsymbol{\xi} \in \mathbb{R}^2$ is the state vector, and η is the parameter that introduces the oscillator nonlinearity. In the next we pick $\eta = 0.1$, $T_s = 0.1$ s, $N_{T_s} = 10$, and $\boldsymbol{\xi}_0 = [1,1]\cdot 10^{-3}$. The sufficient condition C2 in (Aeyels, 1981) yields $N \geq 5$. The changes of V_k in (3) with respect to N is shown in Figure 1. The cost function V_k is shown to become steeper near the local minimum as N increases. Roughly speaking, the higher N, the more positive terms are added in (4). We show the initial V_k, i.e. V_N, since $k \geq N$ (recall that N is the minimum number of sub-samples to fill the window buffer is N). Note that, in general, steeper V_k ensures faster convergence properties for numerical optimisation algorithms.

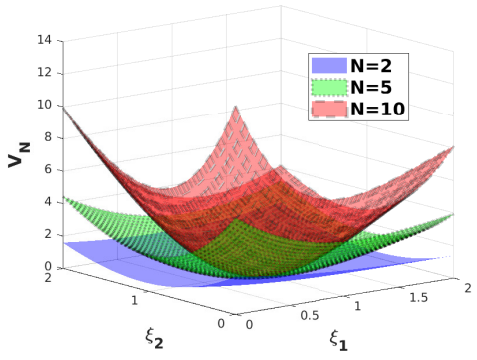

Fig. 1. This figure compares the cost function V_N at increasing values for N.

3.2 Output filtering

Increasing the buffer length N unavoidably yields a computational cost increase. In this section, we propose decreasing N by adding extra measurements via output filtering. Generally speaking, filtered outputs preserve the system's observability under some general assumptions. Results in this sense have been shown in (Menini et al., 2019, 2022). Consider the system (7), with an additional output, namely $\boldsymbol{y} = [\xi_1, \xi_1^f]$, where ξ_1^f comes from a dirty derivative filter (applied to ξ_1) with the following transfer function:

$$G_f(s) = \frac{s}{s\epsilon + 1}, \quad \epsilon \gg 1. \tag{8}$$

As a matter of fact, since we doubled the available outputs, the number of samples can now be chosen with a less strict condition, namely $2N \geq 2n+1$. Therefore, the sufficient condition from (Aeyels, 1981) is relaxed becoming

$$N \geq \left\lceil \frac{2n+1}{p} \right\rceil = \left\lceil \frac{2n+1}{2} \right\rceil, \tag{9}$$

where $\lceil\cdot\rceil$ is an operator returning the ceiling of a number and $p \in \mathbb{N}$ is the number of output signals (including the filter). Clearly, the more the filters, the less strict the condition on the buffer length. Generally speaking, if we consider a number m of filters with different spectral response (with respect to spectral components of $\boldsymbol{y}$), condition (9) turns into

$$N \geq \left\lceil \frac{2n+1}{(1+m)p} \right\rceil, \tag{10}$$

where the notation follows (9). A similar analysis to the one presented in Sec. 3.1 is now carried on the system with the additional output. Again, numeric representation of V_N on the oscillator are reported in Figure 2 picking $N = 5$ when $y = \xi_1$ and both $N = 3$ and $N = 2$ when $\overline{\boldsymbol{y}} = [\xi_1, \xi_1^f]$. The introduction of the filtered signal ξ_1^f makes the cost function steeper near the local minimum $\boldsymbol{\xi}_1$ ($N = 3$ in Figure 2). Note that, if condition (9) does not hold, the cost function in Figure 2 worsens w.r.t. $N = 5$ in Figure 2 with no filters, whereas if even $C2$ fails (necessary condition) then V_N has no longer an isolated minimum in $\boldsymbol{\xi}_1$. Note that discrete-time filter implementation does not add a high computational cost. This algorithm will be referred to as *Filtered* MHE. Due to space constraints, further details on its implementation are given in (Oliva and Carnevale, 2022).

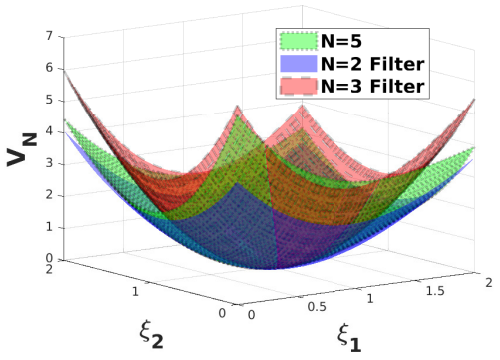

Fig. 2. V_N comparison: no filter with $N = 5$, $N = 3$ and $N = 2$ with filter $G_f(s)$.

3.3 Down-sampling N_{T_s}

The goal of this section is to investigate the role of parameter N_{T_s}, namely the down-sampling factor of the algorithm. If $N_{T_s} = 1$ there is no down-sampling, meaning that the distance between $\boldsymbol{y}_k$ and $\boldsymbol{y}_{k+1}$ is exactly the integration step T_s. This is not a strict condition, as recalled

in (Kang, 2006). However, an incorrect choice of N_{T_s} can significantly deteriorate the performance of the observer, given that the SNR can drastically decrease, depending also on the system trajectory, ruining the performances of the optimization algorithm. Consider the following nonlinear system describing the dynamics of the energy (ξ_1) and anisotropy (ξ_2) of the plasma waves interacting with runaway electrons given by (Buratti et al., 2021):

$$\dot{\xi}_1 = -2\xi_1\xi_2 - 2S + Q, \tag{11a}$$

$$\dot{\xi}_2 = -\nu\xi_2 + \xi_3(\xi_1\xi_2 + S) - \gamma_1 \frac{\xi_2}{1 + \frac{\xi_2}{W_t}}, \tag{11b}$$

$$\dot{\xi}_3 = 0, \tag{11c}$$

$$y = \xi_2, \tag{11d}$$

where Q resembles the energy source provided by plasma interaction, S is the loss of energy due to runaway interaction, $W_t > 0$ is a normalisation factor, and ξ_3 is a parameter that models the interconnection between ξ_1 and ξ_2 dynamics. In this configuration, we are using the proposed observer to estimate the parameter ξ_3 as well (adaptive observer). The rest of the model parameters have been selected as $[Q, S, \gamma_1, \nu, W_t] = [1, 0, 2.5, 0.5, 0.1]$ with $T_s = 0.001$ s , $N = 7 \geq 2n + 1$, with the local optimum for V_N being $\boldsymbol{\xi}_1 = [9.9, 0, 0.5]$, i.e. the initial state of (11) is $\boldsymbol{\xi}_1$. N is picked in order to satisfy C2 and no output filtering was introduced, namely a *Standard* MHE was considered. In order to highlight how the choice of N_{T_s} can strongly affect the observer performance, the observer has been tested with both $N_{T_s} = 5$ and $N_{T_s} = 38$. Numerical simulations show that the observer could correctly reconstruct the state only with $N_{T_s} = 5$. We introduce the following index:

$$\delta_k = \sum_{i=1}^{N-1} \left\| \mathrm{Y}_k^{i+1} - \mathrm{Y}_k^i \right\| + \left\| \boldsymbol{y}_{k+qT_s} - \mathrm{Y}_k^N \right\|. \tag{12}$$

We propose such an index to represent the output signal richness, as it tracks how much the measurements change over time. Note that δ_k is linked to the current measure buffer Y_k, except for the term $\left\| \boldsymbol{y}_{k+qT_s} - \mathrm{Y}_k^N \right\|$, which depends on the current measure $\boldsymbol{y}_{k+qT_s}$. In fact, the index δ_k is computed on the arrival of each new sample, namely every T_s. However, accordingly to the *Standard* MHE, optimisation is launched only every N_{T_s} time instants. This is shown in Figure 3, where the dotted and solid lines show δ_k on each sample, while the squared and circled values correspond to the optimisation. When $N_{T_s} = 38$, the signal richness of the buffer provided to the observer is always nearly zero (Nyquist-Shannon's sampling theorem is unsatisfied), compared to the $N_{T_s} = 5$ case. Intuitively, δ_k provides information on measurement richness reducing the probability of numerical issues, i.e. increases the effectiveness of the minimisation algorithms in solving problem (3).

3.4 Adaptive sampling

As discussed in Sec. 3.3, a time-varying down-sampling can improve the observer's performance. We now propose an adaptive sampling policy under which the observer autonomously selects the best sampling times to fill the buffer. The proposed policy depends on both the signal richness δ_k and the estimation error d_k defined as

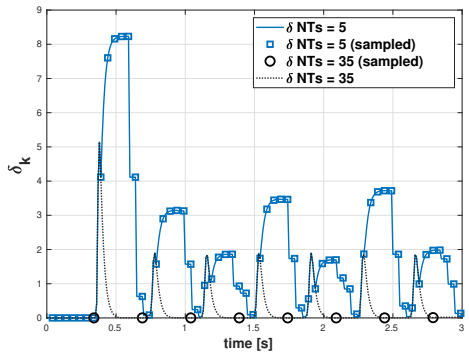

Fig. 3. Comparison of the signal richness δ_k of both $N_{T_s} = 38$ and $N_{T_s} = 5$ cases. The black line shows the case in which $N_{T_s} = 38$, the blue $N_{T_s} = 5$. The circles represent the sampled data when $N_{T_s} = 38$ while the squares when $N_{T_s} = 5$.

$$d_\mathrm{k} = \left\| \mathrm{Y}_k - \hat{\mathrm{H}}_k \right\|. \tag{13}$$

Then, the adaptive sampling time instants satisfy the following conditions

$$t_k : t_k - t_{k-1} \in [N_{T_s}, N_{\max}] \wedge (d_k \geq \bar{\epsilon}_d \wedge \delta_k \geq \bar{\epsilon}_\delta). \tag{14}$$

Note that the number of samples in the buffer does not change and is equal to N. The constrain $d_k \geq \bar{\epsilon}_d \wedge \delta_k \geq \bar{\epsilon}_\delta$ allows to weight the estimation error and the signal richness whereas the constrain $t_k - t_{k-1} \in [N_{T_s}, N_{\max}]$ ensures a down-sampling which is upper and lower limited in time. Indeed, consider the case in which the estimation is sufficiently good. In this situation, the observer could sample the subsequent measurement after a long time. This would result in a considerable amount of model integrations in the next solution of problem (3). Therefore, we also introduced the upper bound $N_{\max}$ over which the buffer is re-initialised with the last N samples equally spaced by N_{T_s}. To summarise, this adaptive sampling policy affects the general structure of the *Standard* MHE described in Sec. 2.1. More specifically, at any time instant, the constraints in (14) are computed to decide whether or not to proceed with the down-sampling and launch the estimation procedure. This algorithm will be referred to as *Adaptive* MHE: details on the algorithm implementation are provided in (Oliva and Carnevale, 2022). Clearly the *Filtered* MHE and the *Adaptive* MHE can be used simultaneously. Simulation results on the adaptive sampling policy are presented in Sec. 4.2.

4. SIMULATION RESULTS

This section presents simulation results of the observer, [2] in order to highlight the performance improvement of the output filtering (see Sec. 3.2) and adaptive sampling (see Sec. 3.4).

4.1 Output filtering

This section focuses on testing the output filtering approach proposed in Sec. 3.2 with a double pendulum model as a test bench. The double pendulum is a widely used benchmark model in control applications (e.g. Shah and

[2] All the simulations have been done on an *AMD Ryzen 5 3500U* with *Radeon Vega Mobile Gfx* with 8 GB RAM.

Rattan (2016)). The general structure of the equations resembles the usual model used in robotics, namely

$$\mathbb{M}(q)\ddot{q} + \mathbb{V}(q,\dot{q})\dot{q} + \mathbb{G}(q) = \tau, \tag{15}$$

where $q = (\theta_1, \theta_2)$ are the angular positions of the links, the $\mathbb{M}$ term describes the system inertia, the $\mathbb{V}(q,\dot{q})$ term the friction and Coriolis effects, and $\mathbb{G}(q)$ the gravitational force. For a more detailed description of the system parameters refer to (Oliva and Carnevale, 2022). The system state is $\boldsymbol{\xi} = [q_1, q_2, \dot{q}_1, \dot{q}_2] \in \mathbb{R}^4$, while the output measurement is $y = q_1 \in \mathbb{R}$. In the forthcoming analysis $T_s = 0.05$ s, and $N_{T_s} = 3$. The output filtering action has been tested by adding to the measurements, both the dirty derivative described in 8 (with $\epsilon = 0.0001$), and another filter, namely an integrator with loss in the following form:

$$G_i(s) = \frac{1}{s\frac{1}{\epsilon_i} + 1}, \tag{16}$$

with $\epsilon_i = 0.1$. In Table 1 the computational times of the observer are reported depending on the number of filters used and on the maximum number of iterations allowed during optimisation. The buffer length was chosen accordingly to (10). All the values reported are scaled with respect to the maximum one, occurring when no filters are used ($N = 9$), and the number of iterations is set to $K = 80$.

Filter n. P	$K = 20$	$K = 40$	$K = 80$
$P = 0$	28%	58%	100%
$P = 1$ (G_f)	14%	29%	50%
$P = 2$ (G_i and G_f)	7.5%	13%	24%

Table 1. *Standard* MHE computation times compared (in percentage) with *Filtered* MHE using either one or two filters.

Therefore, when $P = 0$, the *Standard* MHE algorithm was used, while *Filtered* was used when $P > 0$. In terms of computational speed, output filtering saves a consistent amount of time. Roughly speaking, each filter results in the computational time being halved (check Table 1 columns). The histogram with the distribution of the computational time for each iteration is shown in Figure 4 for all three configurations ($P = 0, P = 1, P = 2$) when $K = 40$. All the data sets have been fitted with a Gaussian distribution. The distributions' characteristics are:

- $N = 9:\ \mu(\mathcal{N}_{N=9}) = 12.4s,\ \sigma(\mathcal{N}_{N=9}) = 1.25s$
- $N = 5:\ \mu(\mathcal{N}_{N=5}) = 5.4s,\ \sigma(\mathcal{N}_{N=5}) = 0.54s$
- $N = 3:\ \mu(\mathcal{N}_{N=5}) = 2.3s,\ \sigma(\mathcal{N}_{N=3}) = 0.14s$

The error dynamics are comparable in all cases, ensuring a good state vector estimation. To better highlight the error dynamics, the norm of the estimation error is reported in logarithmic scale in Figure 5, for the *Standard* MHE and the two *Filtered* MHE, again with $K = 40$. Note that the error starts decreasing after the measurement buffer has been filled. Therefore, the bigger N, the later the optimiser starts the estimation.

4.2 Adaptive sampling results

This section focuses on testing the adaptive sampling policy proposed in Sec. 3.4. The system considered is again (11), with the same setup considered in Sec. 3.4. The maximum number of iterations per optimisation was

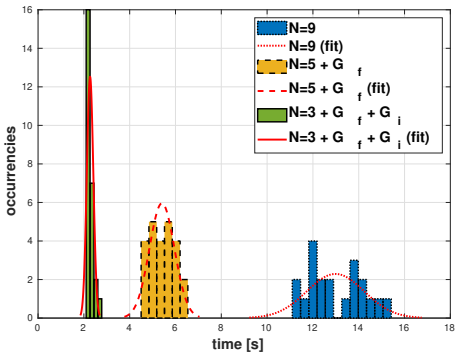

Fig. 4. This figure shows the histogram of the computation time for each iteration of the algorithm with $K = 40$, with $N = 9$ (blue dotted), $N = 5$ (yellow dashed), and $N = 3$ (gree solid). The Gaussian distributions is highlighted following the same line code of the bars.

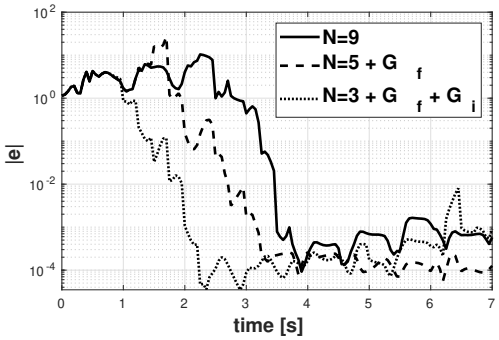

Fig. 5. This figure shows the norm of estimation error on system (15) when the output filtering is used to reduce the minimum number of samples N. The solid line refers to the $N = 9$ configuration, the dashed to $N = 5$, and the dotted to $N = 3$. The error dynamics is shown on a logarithmic scale. In all the reported simulations, $K = 40$ is used.

set to $K = 80$. The *Adaptive* MHE was tested against the *Standard* MHE with a fixed down-sampling factor $N_{T_s} = 5$. Results of this approach are presented in Figure 6 where the estimation error has been reported on a logarithmic scale for both $N = 5$ standard and $N = 5$ adaptive configurations. Please refer to (Oliva and Carnevale, 2022) for a more detailed analysis of the actual sample choice performed by *Adaptive* MHE. The *Adaptive* MHE performance is comparable with the *Standard* MHE, yet slightly worse since the estimation error is thresholded by $\bar{\epsilon}_d$. Clearly, by reducing such value, the estimation accuracy increases. When $\bar{\epsilon}_d = 0$, the *Standard* MHE is obtained again. In terms of computational speed, the adaptive approach saves a consistent amount of time due to a reduced number of optimisation performed. In the presented test case the total computation time decreases from $\mathrm{T}_{tot} = 1130.3$ s with fixed $N_{T_s} = 5$ to $\mathrm{T}_{tot} = 120.9$ s when the adaptive sampling is in use. Again, both the data sets have been fitted with a Gaussian distribution. With the fixed sampling we found $\mu(\mathcal{N}_{\mathrm{fix}}) = 12.9s$ while in the adaptive case, we found $\mu(\mathcal{N}_{\mathrm{adapt}}) = 14.84s$. This is coherent with the expected behaviour, as the adaptive policy does not reduce the single optimisation time but rather the overall number of optimisations. For this reason,

its exploitation for online implementation of the algorithm makes sense in combination with the *Filtered* MHE. For a simulation example of the combined action of *Filtered* MHE and *Adaptive* MHE, refer to (Oliva and Carnevale, 2022).

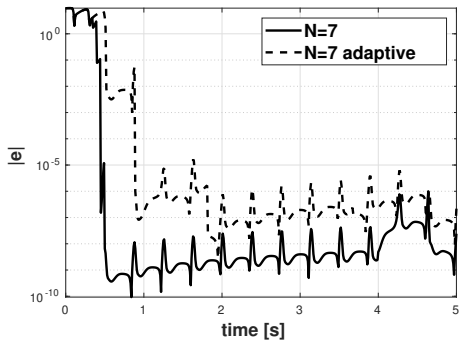

Fig. 6. This figure shows the estimation error on both the Single Shoot (solid line) and adaptive sampling (dashed line) cases on system (11) estimation.

5. CONCLUSIONS AND FUTURE WORK

In the present work, we have initially recalled the Moving Horizon Estimator focusing on the *Standard* MHE, introduced in Sec. 2.1. Following a preliminary analysis, in Sec.3, we introduced two strategies aiming to decrease the computational time of the observer implementation. The critical element for computational improvement stands in adding filtered output. This resulted in the *Filtered* MHE algorithm. The second element consists of implementing an adaptive sampling policy, allowing the observer to autonomously select the *best* measurements to be sampled. This resulted in the *Adaptive* MHE algorithm. Simulation results are provided in Sec. 4 showing a noticeable decrease of computational times up to an order with respect to *Standard* MHE. Future developments of this tool will consider a detailed analysis of the effect of measurement noise and model uncertainties on the observer performance and selection of filters.

REFERENCES

Aeyels, D. (1981). On the number of samples necessary to achieve observability. *Systems & Control Letters*.

Broyden, C. (2000). On the discovery of the "good broyden" method. *Mathematical Programming*.

Buratti, P., Bin, W., Cardinali, A., Carnevale, D., Castaldo, C., D'Arcangelo, O., Napoli, F., Ravera, G.L., Selce, A., Panaccione, L., and et al. (2021). Fast dynamics of radiofrequency emission in ftu plasmas with runaway electrons. *Plasma Physics and Controlled Fusion*.

Carnevale, D., Buratti, P., Bin, W., Bombarda, F., Boncagni, L., Duval, B., Esposito, B., Ceccuzzi, S., Calacci, L., Baruzzo, M., Cappelli, M., Castaldo, C., Centioli, C., Cianfarani, C., Coda, S., Cordella, F., D'Arcangelo, O., Decker, J., Gabellieri, L., Galperti, C., Galeani, S., Garavaglia, S., Ghillardi, G., Granucci, G., Lehnen, M., Liuzza, D., Martinelli, F., Mazzotta, C., Napoli, F., Nardon, E., Oliva, F., Panaccione, L., Passeri, M., Paz-Soldan, C., Possieri, C., Pucella, G., Ramogida, G., Romano, A., Sassano, M., Sheikh, U.A., Tudisco, O., the FTU Team, and the EUROfusion MST1 Team (2021). Results on quiescent and post-disruption runaway electrons studies at frascati tokamak upgrade: RE mitigation via solid deuterium pellets and anomalous doppler instability. *Nuclear Fusion*.

Davis, J.H. (2002). *Luenberger Observers*, 245–254. Birkhäuser Boston.

Glad, S.T. (1983). Observability and nonlinear dead beat observers. In *The 22nd IEEE Conference on Decision and Control*.

Kang, W. (2006). Moving horizon numerical observers of nonlinear control systems. *IEEE Transactions on Automatic Control*.

Karagiannis, D. and Astolfi, A. (2005). Nonlinear observer design using invariant manifolds and applications. In *Proceedings of the 44th IEEE Conference on Decision and Control*.

Karagiannis, D., Carnevale, D., and Astolfi, A. (2008). Invariant manifold based reduced-order observer design for nonlinear systems. *IEEE Transactions on Automatic Control*.

Krener, A.J. and Isidori, A. (1983). Linearization by output injection and nonlinear observers. *Systems & Control Letters*.

Kühl, P., Diehl, M., Kraus, T., Schlöder, J.P., and Bock, H.G. (2011). A real-time algorithm for moving horizon state and parameter estimation. *Computers Chemical Engineering*.

Luders, G. and Narendra, K. (1974). A new canonical form for an adaptive observer. *IEEE Transactions on Automatic Control*.

Marine, R., Santosuosso, G., and Tomei, P. (2001). Robust adaptive observers for nonlinear systems with bounded disturbances. *IEEE Transactions on Automatic Control*.

Marino, R. (1990). Adaptive observers for single output nonlinear systems. *IEEE Transactions on Automatic Control*.

Marino, R. and Tomei, P. (1992). Global adaptive observers for nonlinear systems via filtered transformations. *IEEE Transactions on Automatic Control*.

Menini, L., Possieri, C., and Tornambe, A. (2022). On the use of the time-integrals of the output in observer design for nonlinear autonomous systems. *IEEE Transactions on Automatic Control*.

Menini, L., Possieri, C., and Tornambè, A. (2019). Observers for linear systems by the time integrals and moving average of the output. *IEEE Transactions on Automatic Control*.

Michalska, H. and Mayne, D. (1995). Moving horizon observers and observer-based control. *IEEE Transactions on Automatic Control*.

Oliva, F. and Carnevale, D. (2022). On the implementation of adaptive and filtered mhe. *Arxiv*.

P.E. Moraal, J.G. (1995). Observer design for noninear systems with discrete-time measurements. *IEEE Transactions on Automatic Control*.

Reif, K., Gunther, S., Yaz, E., and Unbehauen, R. (1999). Stochastic stability of the discrete-time extended kalman filter. *IEEE Transactions on Automatic Control*.

Schiller, J.D., Knüfer, S., and Müller, M.A. (2021). Robust stability of suboptimal moving horizon estimation using an observer-based candidate solution. *IFAC-PapersOnLine*.

Shah, J. and Rattan, S. (2016). Dynamic analysis of two link robot manipulator for control design using pid computed torque control. *IAES International Journal of Robotics and Automation (IJRA)*.

Sui, D., Johansen, T.A., and Feng, L. (2010). Linear moving horizon estimation with pre-estimating observer. *IEEE Transactions on Automatic Control*.

Suwantong, R., Bertrand, S., Dumur, D., and Beauvois, D. (2014). Stability of a nonlinear moving horizon estimator with pre-estimation. In *2014 American Control Conference*.

Thrun, S., Burgard, W., and Fox, D. (2006). *Probabilistic Robotics*. The MIT Press.

Tousain, R., van der Meche, E., and Bosgra, O. (2001). Design strategy for iterative learning control based on optimal control. *Proceedings of the 40th IEEE Conference on Decision and Control (Cat. No.01CH37228)*. doi:10.1109/CDC.2001.980905.

Tyukin, I.Y., Steur, E., Nijmeijer, H., and van Leeuwen, C. (2013). Adaptive observers and parameter estimation for a class of systems nonlinear in the parameters. *Automatica*.

Wynn, A., Vukov, M., and Diehl, M. (2014). Convergence guarantees for moving horizon estimation based on the real-time iteration scheme. *IEEE Transactions on Automatic Control*.