Research Article

Enhancement of the Moving Horizon Estimation Performance Based on an Adaptive Estimation Algorithm

Steve Alan Talla Ouambo, Alexandre Teplaira Boum, Adolphe Moukengue Imano, and Jean-Pierre Corriou

1University of Douala, Faculty of Science, Douala 24157, Cameroon
2University of Douala, ENSET, Douala 1872, Cameroon
3University of Lorraine, LRGP, UMR, Nancy 7274, France

Correspondence should be addressed to Steve Alan Talla Ouambo; tallaouambo_alan@yahoo.fr

Received 1 September 2021; Revised 29 November 2021; Accepted 13 December 2021; Published 29 December 2021

Abstract

Although moving horizon estimation (MHE) is a very efficient technique for estimating parameters and states of constrained dynamical systems, however, the approximation of the arrival cost remains a major challenge and therefore a popular research topic. The importance of the arrival cost is such that it allows information from past measurements to be introduced into current estimates. In this paper, using an adaptive estimation algorithm, we approximate and update the parameters of the arrival cost of the moving horizon estimator. The proposed method is based on the least-squares algorithm but includes a variable forgetting factor which is based on the constant information principle and a dead zone which ensures robustness. We show by this method that a fairly good approximation of the arrival cost guarantees the convergence and stability of estimates. Some simulations are made to show and demonstrate the effectiveness of the proposed method and to compare it with the classical MHE.

1. Introduction

Since the 1960s, the state estimation of systems by a moving horizon approach has been highlighted. The dazzling success of this estimation technique is based on its ability to explicitly take into account the constraints on the estimate of variables of a dynamic system [1–3]. In [4], the problem of estimating the state of a linear system on a peer-to-peer network of linear sensors is discussed. The proposed approach is fully distributed and scalable and allows for taking into account the constraints on noise and state variables by resorting to the moving horizon estimation paradigm. Each node in the network calculates its local state estimate by minimizing a cost function defined over a sliding window of fixed size. The cost function includes a fused arrival cost which is computed in a distributed way by performing a consensus on the local arrival costs. In [5], the authors have proposed a novel type of iterative partition-based moving horizon estimators, the estimates of which approach those of a centralized moving horizon estimator as the number of iterations increases. It uses a deterministic setting and can handle known inputs as well as bounds on the estimated state. They have derived conditions on the system, its partitions, and the scalar regularization parameter, which guarantee convergence towards the optimal centralized state estimate as well as stability of the estimation error dynamics, even with a finite number of iterations at each time step. The authors, in [6], have investigated on a problem of event-triggered state estimation for networked linear systems. They consider that the stochastic system disturbances and noise are bounded and moving horizon estimation (MHE) is used to handle these constraints and establish an event-based state estimation mechanism that aims to provide good state estimates while reducing the frequencies of both the evaluation of the state estimator and networked communication between the plant and the estimator. In [7], the authors focus on distributed moving horizon estimation (DMHE) for a class of two-time-scale nonlinear systems described in the
framework of singularly perturbed systems. By taking advantage of the time-scale separation property, a two-time-scale system is first decomposed into a reduced-order fast system and a reduced-order slow system. The slow system is further decomposed into several interconnected slow subsystems. In the proposed distributed state estimation scheme, a local estimator is designed for each slow subsystem and for the reduced-order fast system.

In the moving horizon estimation approach, the state estimate is determined by solving an optimization problem online that minimizes the sum of the squared errors. At a sampling time, when a new measurement is obtainable, the older measuring within the estimation window is discarded and the horizon finite optimization problem is solved again to obtain the new state estimate [8–10]. To take into account the historical data outside the estimation window, a variable that summarizes the information from this data is included in the objective function of the MHE optimization problem called arrival cost. This arrival cost is crucial for the stability and performance of the moving horizon estimation; therefore, its approximation is an active research topic.

To reduce the size of the estimation window and the size of the optimization problem, a good approximation of the arrival cost is necessary, ensuring good performance and good robustness. The arrival cost is commonly estimated using a weighted norm of states at the beginning of the estimated horizon window [8, 11, 12]. In [13], the authors proposed to use the Kalman filter approach for updating the term of the arrival cost (in particular the weighting matrix) for the case of linear systems, leading to an inaccurate approximation of the weighting matrix leading to relatively poor estimates due to a Gaussian distribution of the Kalman filter. To overcome this problem, the authors in [14] proposed an iterative scheme for updating the arrival cost using the information on the active or inactive constraints from the previous iteration and a quadratic approximation. The idea brings the quadratic approximation as close as possible to the optimal solution of the arrival cost. The fundamental assumption behind the updating scheme is that the constraints should remain unchanged after a certain moment despite the increase of the estimation horizon. For a large estimation window, this assumption is functional giving good results; however, if after smoothing some constraints become inactive, less good results are obtained.

In [15], the author uses nonlinear filters based on sampling, including the unscented Kalman filter based on deterministic sampling, the particle filter based on random sampling, and the cell filter based on the aggregate Markov chain to update the arrival cost of the moving horizon estimation and demonstrate the advantages and performance of those filters over the traditional extended Kalman filter approach. However, the computation time is relatively long and better performance is also obtained due to a fairly large horizon length.

In more recent works, as in [16], the authors proposed a moving horizon estimation algorithm based on multiple estimation windows. The major advantage of this algorithm is that it reduces the size of the optimization problem while maintaining the performance properties of the estimator and its stability by taking advantage of the inactivity of the constraints. In [17], the authors developed a two-step design strategy, namely, decoupling step and convergence step for networked linear systems with unknown inputs under dynamic quantization effects. In the decoupling step, the decoupling parameter of the moving horizon estimator is designed based on certain assumptions on system parameters and quantization parameters. In the convergence step, by employing a special observability decomposition scheme, the convergence parameters of the moving horizon estimator are achieved such that the estimation error dynamics is ultimately bounded. The conventional moving horizon estimation strategy turns out to be unable to guarantee satisfactory performance as the estimation error depends on external disturbances. In [18], a moving horizon estimation strategy is formulated and the corresponding analytical solution of the state estimation is derived by using the completing the square technique. Based on the lifting approach, the dynamics of the estimation error covariance is analyzed and the desired estimator parameters are obtained by solving a set of linear matrix inequalities.

In this work, the weighting matrix is updated using an adaptive estimation algorithm, which is the least-squares algorithm including a dead zone and a variable forgetting factor [19, 20]. The dead zone here is used to obtain the robustness; the formula of a forgetting factor based on the principle of constant information can be applied to the process model with unmodeled process uncertainties or external disturbances. Uncertainties can cause the estimates to differ when the standard recursive least-squares algorithm is used, but to prevent this, the inclusion of the dead zone in the estimation algorithm is required. Therefore, the weighting matrix will be calculated in closed loop unlike standard estimation techniques. The effectiveness of the proposed approach is evaluated by carrying out simulation studies on benchmark problems found in the literature.

The main contributions of this work aim to (i) show that an adaptive estimation method can be implemented to approximate the arrival cost and have much better efficiency than the traditional MHE, (ii) also show that the proposed scheme guarantees robustness to the estimator for sudden disturbances, and (iii) finally show that the bounds of estimates are crucial when updating the arrival cost. The paper is organized as follows. In Section 2, we present the moving horizon estimation problem, followed by the arrival cost update mechanism in Section 3. Section 4 presents a stability analysis; the results are discussed in Section 5, and finally, Section 6 will be devoted to the conclusions.

2. Preliminaries and Problem Statement

Notation 1. In the following, we will consider the following notations, \( \xi \in \mathbb{R}^n \) is a column vector and \( \xi^T \in \mathbb{R}^n \) will be its transpose. \( \xi_{jt} \) denotes a finite sequence of vectors over a given index up to \( t \). Let \( B \in \mathbb{R}^{n \times n} \) be a \( n \times n \) matrix and \( B^T \) is its transpose. If \( B \in \mathbb{R}^{n \times n} \), then \( B^{-1} \) denotes its inverse. Let \( P \in \mathbb{R}^{n \times n} \) be a symmetric real matrix, \( P \) is said to be positive definite if \( \xi^T P \xi > 0 \) and \( \forall \xi \in \mathbb{R}^n \). The Euclidean matrix or
vector norm is denoted by \( \| \cdot \| \), for instance, \( \xi \in \mathbb{R}^n \) and \( P \in \mathbb{R}^{n \times n} \) are positive definite; it follows that \( \| \xi \|_P = \sqrt{\xi^T P \xi} \).

Consider the linear time invariant discrete system:

\[
x_{t+1} = Ax_t + Gw_t, \\
y_t = Cx_t + v_t,
\]

where \( x_t \in X_0 \subseteq \mathbb{R}^n \) is the state vector, \( w_t \in W_0 \subseteq \mathbb{R}^p \) is the process noise vector, \( y_t \in Y_0 \subseteq \mathbb{R}^m \) is the measurement vector, and \( v_t \in V_0 \subseteq \mathbb{R}^m \) is the measurement noise vector. The process noise \( w_t \) and the measurement noise \( v_t \) are assumed to be bounded and unknown eventually, i.e., \( w \in W_0(\theta_{W_0\text{max}}) \) and \( v \in V_0(\theta_{V_0\text{max}}) \) for \( (\theta_{W_0\text{max}}, \theta_{V_0\text{max}}) \in \mathbb{R}_+^2 \), with \( X_0, Y_0, W_0, \) and \( V_0 \) being compact and convex sets considered closed with \( 0 \in X_0, 0 \in Y_0, 0 \in W_0, \) and \( 0 \in V_0 \).

Using all measures for the full information filter to estimate the system states \( \bar{x}_t \) increases the size of the optimization problem over time and leads to heavy calculation.

To overcome this problem, the moving horizon estimation was developed to take into account a certain amount of data and possibly to make dynamic updates of certain parameters in an estimation window moving in time. A penalty term called arrival cost is essential in the formulation of the MHE because it takes into account past information into the problem, in case of which a poor estimate will be revealed.

In many dynamical systems, in particular linear systems, a Gaussian distribution is assumed and the arrival cost function is updated using a Kalman filter [23]. Besides, the MHE is equivalent to the Kalman filter for unconstrained linear systems with indeed the same Gaussian distribution. In the following, we will present the proposed mechanism for updating the arrival cost parameters of the MHE.

### 3. Arrival Cost Update Algorithm

For the constrained MHE problem, it is difficult to find an exact analytical expression of the arrival cost; therefore, an approximation of the arrival cost used for the unconstrained problem will be used; this quadratic approximation is of the following form:

\[
\Gamma_{\ell-N} = \| \bar{x}_{\ell-N} - \bar{x}_{\ell-1} \|_{P_{\ell-N}^{-1}}^2,
\]

where \( P_{\ell-N} \) is the weighting matrix and \( \bar{x}_{\ell-N} \) is the initial state. Both define the approximation of the arrival cost function. The proposed approach for updating the arrival cost in this work, based on an adaptive algorithm, is formulated as an update of the following elements: \( \bar{x}_{\ell-N} \) and \( P_{\ell-N} \).

The update of the initial state \( \bar{x}_{\ell-N} \) will be made by a smooth update, which implies that once the estimate leaves the estimation horizon, the state of the constraints does not change [14, 24]:

\[
\bar{x}_{\ell-N} = \bar{x}_{\ell-N|\ell-1}.
\]

Regarding the update of the weighting matrix \( P_{\ell-N} \), it is necessary to notice from the outset that, in a stochastic interpretation of the MHE problem, the weighting matrix can be seen as the covariance of the state \( x_{\ell-N} \) [8]. For unconstrained linear systems, analytical solutions exist to compute the weighting matrix; on the contrary, for constrained linear systems, a deduction of the approximate solutions is made, which are fundamentally and based on a recursive process of updating useful information from the system at a given time. In the proposed adaptive estimation method, an update based on a recursive process relying on states (or not) and measurements is used [19, 25]:

\[
P_{\ell-N}^{-1} = \lambda_{\ell} P_{\ell-N-1}^{-1} + \frac{\alpha_{\ell}}{\lambda_{\ell}} \bar{x}_{\ell-N|\ell-1} \bar{x}_{\ell-N|\ell-1}^T.
\]

With \( P_0 > 0 \), where \( \lambda_{\ell} \) is the forgetting factor with \( 0 < \lambda_{\ell} \leq 1 \), \( \alpha_{\ell} \) is a function which varies between 0 and 1 called the dead zone. Equation (6) produces a recursively
real-time estimate of $P_{\ell-N}$, by updating $P_{\ell-N-1}$ (previous estimate) with a dynamic temporal average over time of matrix $\tilde{x}_{\ell-N|\ell-1}^T \tilde{x}_{\ell-N|\ell-1}$. We can qualify this updating mechanism as a dynamic temporal filter overtime with $P_0$, the initial condition and $\tilde{x}_{\ell-N|\ell-1} \tilde{x}_{\ell-N|\ell-1}$ are the input.

**Lemma 1.** If a matrix $Z > 0$, then there exists an invertible matrix $A$ such that $A^TZA = I$, which is the identity matrix. Moreover, $AA^T = Z^{-1}$ which exists.

**Proof.** See Lemma 7 in [26].

Note that the forgetting factor $\lambda_\ell$ and the function $\alpha_\ell$ each have a different role. When $\lambda_\ell \approx 1$, the algorithm tends to keep the old data in $P_{\ell-N-1}$; in this case, all information is included, improving the estimation problem. On the contrary, when $\lambda_\ell \ll 1$, the algorithm tends to discard the old information in $P_{\ell-N-1}$, considering recent data to estimate $P_{\ell-N}$ [27]. This case allows avoiding many old data that could negatively affect the estimator performance. The dead zone function $\alpha_\ell$ is used to ensure robustness when estimating.

By the matrix inversion Lemma, we can rewrite (6) as follows:

$$P_{\ell-N} = \frac{1}{\lambda_\ell} \left[ P_{\ell-N-1} - \alpha_\ell P_{\ell-N-1} \tilde{x}_{\ell-N|\ell-1}^T \tilde{x}_{\ell-N|\ell-1} P_{\ell-N-1} \right].$$

A feedback loop is introduced into the MHE problem by this way to compute $P_{\ell-N}$; this way makes it possible to ensure the convergence of the estimates $\tilde{x}_{\ell-N|\ell}$ and $\tilde{w}_j (j = \ell - N, \ldots, N)$ and to adjust $P_{\ell-N}$ according to the amount of available information. The adaptive algorithm used proposes a variable forgetting factor varying between 0 and 1 and under certain conditions to be respected and a dead zone, both allowing $P_{\ell-N}$ to be updated as follows [19]:

$$\lambda_\ell = \begin{cases} \frac{(N_0 + \Delta^2_\ell)}{(N_0 + \Delta^2_\ell + M_{1\ell})} & \text{if } \Delta^2_\ell < 1, \\ 1, & \text{otherwise,} \end{cases}$$

$$\alpha_\ell = \begin{cases} 1, & \text{if } \Delta^2_\ell \left( 1 + \frac{\mu_\ell}{\lambda_\ell} \right) \leq \epsilon^2_\ell, \\ 0, & \text{otherwise,} \end{cases}$$

where $N_0$ is a positive design parameter chosen which is quite large: $M_{1\ell} = e_\ell^2/1 + \mu_\ell$, $\mu_\ell = \tilde{x}_{\ell-N|\ell-1}^T P_{\ell-N-1} \tilde{x}_{\ell-N|\ell-1}$, and $e_\ell = y_{\ell-N} - \tilde{y}_{\ell-N}$, where $\Delta_\ell = \|\tilde{x}_{\ell-N|\ell-1}\|d_1 + d_2/n_{\ell-1}$, with $n_{\ell-1} = \max (1, \|\tilde{x}_{\ell-N|\ell-1}\|)$ and $d_1$ and $d_2$ are also design parameters, which are chosen positive.

Any sequence $\{\alpha_\ell\}$ is bounded by some $\sigma^* > 0$. Thus, the size of dead zone $\Delta^2_\ell (1 + \mu_\ell/\lambda_\ell)$ is also bounded because $\mu_\ell$ is bounded too.

The forgetting factor proposed in [19] is based on the constant information principle, which is adequate for external disturbances and unmodeled system uncertainties.

The following algorithm summarizes the proposed moving horizon estimation estimator based on the adaptive estimation method for the arrival cost update.

**Algorithm 1.** $\tilde{x}_0$, $P_0$, $Q, R > 0$.

The stability of the estimator requires that the prediction error converges to zero for the following nominal system:

$$x_{\ell+1} = Ax_\ell,$$

$$y_\ell = Cx_\ell,$$

with no measurement and state noise. In [8], the authors require that the system has to evolve according to the constraints because, for the constrained systems’ case, a poor range of choice of constraints can prevent convergence to the real values of the system.

**Lemma 2.** The estimation algorithm assumes that $\sup (\|e_\ell\|) \leq d_1$ and $\sup (\|w\|) \leq d_2$, with $\delta_{\ell}$ as the system uncertainty vector; then, the following properties are valid:

(i) $P_\ell$ converges and is uniformly bounded

(ii) $\sup (e_\ell^2 - (1 + K)\Delta^2_\ell) \leq 0$, $\forall K \geq 0$, with $e_\ell = y_\ell - \tilde{y}_\ell$

**Proof.** By recursive calculation and the matrix inversion lemma, we get from (7)

$$P_\ell^{-1} = P_0^{-1} + \sum_{i=1}^{\ell} \frac{\alpha_i}{\sigma_{i\ell}} \tilde{x}_{i\ell-1} \tilde{x}_{i\ell-1}^T.$$

It follows that

$$\|P_\ell\| \leq \frac{(1 + v_0)}{N_0} \|P_0\|,$$

which implies the uniform boundedness of $P_\ell$; by the property $0 < \sigma^* \leq \sigma_{0,\ell}$ it is easily proven from (11) that $P_\ell$ uniformly converges. □
Remark 1. The matrix $P_\ell$ owns a steady-state solution, i.e.,
\[
\lim_{\ell \to \infty} P_\ell = P_\infty.
\]

Proof. The update expression of $P_{\ell-1}$ given by equation (7) can be seen as a particular case of Ricatti equation:
\[
\lambda_\ell P_{\ell-1} = P_{\ell-1} - \frac{\alpha_\ell P_{\ell-1} \bar{x}_{\ell-1}^T \bar{x}_{\ell-1} P_{\ell-1}}{\lambda_\ell + \bar{x}_{\ell-1}^T \bar{x}_{\ell-1}}
\]
\[
= P_{\ell-1} - \frac{\alpha_\ell P_{\ell-1} \bar{x}_{\ell-1}^T \bar{x}_{\ell-1} P_{\ell-1}}{\lambda_\ell + \bar{x}_{\ell-1}^T \bar{x}_{\ell-1}}.
\]

(14)

It is stable if and only if the eigenvalues of the transfer matrix are strictly inside the unit circle of the complex plane:
\[
\Lambda = I - \frac{\alpha_\ell P_{\ell-1} \bar{x}_{\ell-1}^T \bar{x}_{\ell-1} P_{\ell-1}}{\lambda_\ell + \bar{x}_{\ell-1}^T \bar{x}_{\ell-1}},
\]
\[
= I - \tilde{\Lambda},
\]

(15a)

(15b)

\[
\nu_\ell - \lambda_\ell \nu_{\ell-1} = \bar{x}_{\ell-1}^T P_{\ell-1} \bar{x}_{\ell-1} - \lambda_\ell \bar{x}_{\ell-1}^T P_{\ell-1} \bar{x}_{\ell-1}
\]
\[
= (A\bar{x}_{\ell-1})^T P_{\ell-1} \bar{x}_{\ell-1} - \lambda_\ell \bar{x}_{\ell-1}^T P_{\ell-1} \bar{x}_{\ell-1}
\]
\[
= \bar{x}_{\ell-1}^T A^T P_{\ell-1} A \bar{x}_{\ell-1} - \lambda_\ell \bar{x}_{\ell-1}^T P_{\ell-1} \bar{x}_{\ell-1}
\]
\[
= \bar{x}_{\ell-1}^T A^T \left[ \lambda_\ell \left( P_{\ell-1} + \frac{\alpha_\ell}{\lambda_\ell} \bar{x}_{\ell-1}^T \bar{x}_{\ell-1} \right) \right] A \bar{x}_{\ell-1}
\]
\[
- \lambda_\ell \bar{x}_{\ell-1}^T P_{\ell-1} \bar{x}_{\ell-1}
\]
\[
= \bar{x}_{\ell-1}^T A^T \left[ \lambda_\ell P_{\ell-1} + \frac{\alpha_\ell}{\lambda_\ell} \bar{x}_{\ell-1}^T \bar{x}_{\ell-1} A - P_{\ell-1} \right] \bar{x}_{\ell-1},
\]

(18)

and $A^T \bar{x}_{\ell-1} A$ are symmetric positive definite and symmetric positive semidefinite, respectively, for all $\ell$; then, if (10) is observable, it follows that the MHE with the proposed adaptive arrival cost update is an asymptotically stable observer.

Regarding bounded stability, we can note that the arrival cost satisfies

\[
\lambda_{\min}(P_{\ell-1}^{-1}) \frac{\|\bar{x}_{\ell-1} - \bar{x}_{\ell-1}\|}{\|\bar{x}_{\ell-1} - \bar{x}_{\ell-1}\|}
\]
\[
\leq \|\bar{x}_{\ell-1} - \bar{x}_{\ell-1}\|_{P_{\ell-1}^{-1}}
\]
\[
\leq \lambda_{\max}(P_{\ell-1}^{-1}) \frac{\|\bar{x}_{\ell-1} - \bar{x}_{\ell-1}\|}{\|\bar{x}_{\ell-1} - \bar{x}_{\ell-1}\|}
\]
\[
\cdot \lambda_{\min}(Q^{-1}) \|\tilde{w}\| \leq \|\tilde{w}\|_{Q^{-1}} \leq \lambda_{\max}(Q^{-1}) \|\tilde{w}\|
\]
\[
\cdot \lambda_{\min}(R^{-1}) \|\tilde{v}\| \leq \|\tilde{v}\|_{R^{-1}} \leq \lambda_{\max}(R^{-1}) \|\tilde{v}\|.
\]

(19)
The estimator with adaptive arrival cost update is robust globally asymptotically stable under bounded disturbances; besides, if \( \lim_{t \to \infty} v_t = 0 \) and \( \lim_{t \to \infty} w_t = 0 \), then \( \lim_{t \to \infty} \| x_t - \tilde{x}_t \| = 0 \). □

Proof. cf. Theorem 7 in [28]. □

5. Simulation Results

To demonstrate the performance of the proposed method, we adopted discrete-time models used by [13, 29].

5.1. Example 1.

\[
\begin{align*}
    x_{t+1} &= \begin{bmatrix} 0.99 & 0.2 \\ -0.1 & 0.3 \end{bmatrix} x_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w_t, \\
    y_t &= \begin{bmatrix} 1 & -3 \end{bmatrix} x_t + v_t,
\end{align*}
\]

(20)

where \( w_t = [s_t] \) with \( [s_t] \) a sequence of independent, normally distributed random values with zero mean and covariance equal to 1, and \( [v_t] \) is also a sequence of independent, normally distributed random values with zero mean and covariance equal to 0.1. The initial state \( \bar{x}_0 \) with zero mean is also normally distributed with a unit covariance. The constrained estimation problem here is formulated with \( Q = I, \ R = 0.01, \ P_0 = 0.5I, \) and \( \bar{x}_0 = [0.5, 0.5]^T \). We choose \( N = 10 \) and \( d_1 = 0.01 \) and \( d_2 = 0.005 \) and \( N_0 = 50 \) as design parameters.

The proposed algorithm is compared with the classical MHE (MHEKF) which propagates the initial state and the arrival cost estimate using Kalman filter. The sum square estimation error is used as a benchmark:

\[
x^{(i)} = \sum_{t=0}^{\ell} (x^{(i)}_t - \tilde{x}^{(i)}_t)^2,
\]

(21)

where \( x^{(i)} \) denotes the \( i^{th} \) component of the vector \( x \). Based on 70 trials, the average sum square estimation error was computed. The results are shown in Table 1, where it can be seen that the performance of MHE\(_{AD} \) is superior to the classical MHE.

Table 1: Average sum square estimation error (example 1).

|   | MHE\(_{AD} \) | MHE\(_{KF} \) |
|---|-------------|-------------|
| \( x^{(1)} \) | 44.01       | 83.16       |
| \( x^{(2)} \) | 70.97       | 125.10      |

Some variations between 0 and 1 to allow a rapid adaptation of the weighting matrix when a sudden change occurs. We can note that when the estimation error is large, the dead zone function is equal to 1, and when the error is relatively low, the dead zone function is equal to 0. Figures 3(a) and 3(b) allow appreciating the evolution of the trace of the weighting matrix \( P_t \) for both estimators.

The MHE\(_{KF} \) weighting matrix converges to the steady-state value and does not change anymore, while MHE\(_{AD} \) adapts its weighting matrix according to variations of the process noise.

A particular case is studied and presented in Figure 4, where the process noise covariance \( \sigma_w \) varies with time. The covariance will take values 0.9 when \( 0 \leq t < 5 \), 0.55 when \( 5 \leq t < 20 \), 0.8 when \( 20 \leq t < 40 \), 1.3 when \( 40 \leq t < 60 \), and 0.4 when \( 60 \leq t < 80 \). We can see in Figures 4(a) and 4(b) that the proposed approach still performs efficiently better than the conventional MHE despite the variation with time of the covariance. We can also note that compared to the previous scenario, the estimate is less good for both estimators, but this can be explained by the fact that the estimators have no knowledge of this happening. Figures 4(c) and 4(d) show the trace of the matrix \( P_t \) used by MHE\(_{KF} \) and MHE\(_{AD} \), respectively. We can observe that the MHE\(_{KF} \) weighting matrix converges to a steady value at 20s and does not change anymore, while the proposed method tries to adapt its weighting matrix according to the variations of the process noise.

It can be seen that the trace of \( P_t \) tends to stabilize when the error is small. On the contrary, the trace of the weighting matrix becomes smaller when the difference between the estimate and the real state grows, ensuring good adaptation despite the influence of varying disturbances.

In Figures 4(c) and 4(f), we can see the evolution of the forgetting factor and the dead zone, respectively.

We can see that the evolution of the forgetting factor is independent of the dead zone, but it evolves according to the variation of the process noise. We also notice that the forgetting factor tries to adapt to the variations while keeping its value close to 1. On the contrary, the behavior of the dead zone is such that when the estimation error is small, the dead zone function tends to 0 while when the error grows, the dead zone is equal to 1, so to allow a good adaptation and provide an acceptable estimation.

Figure 5 shows the performance of MHE\(_{AD} \) with different values of the horizon length. We can notice that, for small values of \( N \), the estimation error is not negligible, while for large values of \( N \), the estimation error is very appreciable. Table 2 allows us to assess the results quantitatively.

5.2. Example 2. Still in the same logic of showing the performance of the proposed approach, the example of the following system is provided:
where
\[
\begin{align*}
x_{t+1} &= Ax_t + Bu_t + B_d d_t + w_t, \\
y_t &= Cx_t + v_t,
\end{align*}
\] (22)

Here, disturbances are modeled with the same distribution as example 1. The constrained estimation problem is formulated with \( Q = I \), \( R = \begin{bmatrix} 20 & 0 \\ 0 & 2 \end{bmatrix} \), \( P_0 = 10I \), and \( \overline{x}_0 = [1.5 -0.5 \ 0]^T \).

\( u_t = 4H(\ell - 0.15) - 4H(\ell - 2.5) \) and \( d_t = -37H(\ell - 0.25) + 37H(\ell - 0.7) + 37H(\ell - 1.3) - 37H(\ell - 1.7) - 40H(\ell - 4.5) \) are the control input and the disturbance, respectively. \( H \) represents the Heaviside function. We also choose the horizon \( N = 10 \) in this example and \( d_1 = 0.00005 \), \( d_2 = 0.001 \), and \( N_0 = 5000 \) as design parameters. Figures 6(a)–6(c) show the states and their estimates. From these figures, we can see that the proposed method (MHEAD) gives a good estimate compared to the conventional MHE (MHEKF); this is because, in this example, the inputs of the system are bounded and persistent. We can simply say the MHEAD estimates can track properly the states. The divergence of the conventional MHE is due to the inclusion of a large value of the disturbance (sudden variation) at \( t = 4.5 \) s; on the contrary, the
Figure 2: Evolution of the sequence of (a) forgetting factor ($\lambda_k$) and (b) dead zone function ($\alpha_k$).

Figure 3: Evolution of trace ($P_k$). (a) Trace ($P_k$) – MHEKF. (b) Trace ($P_k$) – MHEAD.

Figure 4: Continued.
The proposed algorithm adapts to slow or sudden changes. This can easily be explained by the fact that the crucial parameters of the proposed algorithm can follow slow or sudden changes or disturbances and provide good adaptability. Figures 7(a) and 7(b) show the evolution of the forgetting factor and the dead zone function, respectively. The same remarks as in the previous example are also considered in this case. It can be seen that the best performance of the proposed method occurred when the forgetting factor was equal to 1, and the dead zone function
Table 2: Comparison of the effect of the horizon size on the average sum square estimation error for states $x^{(1)}$ and $x^{(2)}$.

| $N$  | $x^{(1)}$ | MHE$_{AD}$ | $x^{(2)}$ |
|------|-----------|-------------|-----------|
| 3    | 77.52     | 88.41       |
| 6    | 67.03     | 80.05       |
| 10   | 48.11     | 71.15       |

Figure 6: Continued.
Figure 6: Comparison of estimators for model (22) with horizon length $N = 10$. (a) Estimation of $x^{(2)}$. (b) Estimation of $x^{(2)}$. (c) Estimation of $x^{(3)}$. (d) $(x^{(1)} - \hat{x}^{(1)})^2$. (e) $(x^{(2)} - \hat{x}^{(2)})^2$. (f) $(x^{(3)} - \hat{x}^{(3)})^2$.

Figure 7: Evolution of the sequence of (a) forgetting factor ($\lambda_k$) and (b) dead zone function ($\alpha_k$).

Figure 8: Evolution of trace ($P_k$). (a) Trace ($P_k$) – MHEKF. (b) Trace ($P_k$) – MHEAD.
equal to 0, when the disturbances were less important. When the disturbances were considerable, the forgetting factor tried to adjust to them. Figures 8(a) and 8(b) allow appreciating the evolution of the trace of the weighting matrix $P_\ell$ for both estimators. We can see that the evolution of the trace of the classical MHE weighting matrix has difficulty reaching the steady state unlike the previous example, which justifies the weak reconstruction; on the contrary, the evolution of the weighting matrix of MHEAD is smoother; that is, when the disturbances are less consequent, the evolution of the trace of the weighting matrix of MHEAD behaves as a classical constrained optimization problem. When the noise increases, the proposed estimator does not follow the system and the forgetting factor decreases and

| Table 3: Average sum square estimation error (example 2). |
|-----------------|-----------------|-----------------|
| $x^{(1)}$       | MHEAD           | MHEKF           |
| $x^{(2)}$       | 75.37           | 185.10          |
| $x^{(3)}$       | 115.02          | 201.50          |
|                 | 129.25          | 198.22          |

| Table 4: Comparison of the effect of the horizon size on the average sum square estimation error for states: $x^{(1)}$, $x^{(2)}$, and $x^{(3)}$. |
|-----------------|-----------------|-----------------|
| $N = 3$         | $x^{(1)}$       | 78.01           | 122.02          | 133.41 |
| $N = 6$         | 76.95           | 119.55          | 132.01          |
| $N = 10$        | 75.37           | 115.02          | 129.25          |
moves away from 1 to allow a fairly fast adaptation of $P_\ell$. Based on 70 trials, the average sum square estimation error was computed, and the results are shown in Table 3. Likewise, in Figure 9, we show the performance of $\text{MHE}_\lambda$ with different values of the horizon length. It can be seen that $\text{MHE}_\lambda$ errors do not change significantly with the different horizon sizes, showing that this algorithm seems to be a good method for the arrival cost approximation, even for short horizons. Table 4 allows us to assess the results quantitatively.

6. Conclusion

In this paper, an alternative method to approximate the arrival cost for the moving horizon estimation has been studied on linear systems. The proposed approach is based on the least-squares algorithm but includes a variable forgetting factor which is based on the constant information principle and a dead zone which ensures robustness. The use of CasADi toolbox [30] was used for the implementation of the optimization algorithm for real-time operation. The results obtained show that the adaptive estimation algorithm allows a good approximation of the arrival cost for the moving horizon estimation. This can easily be explained by the fact that the crucial parameters of the proposed method can follow slow or sudden changes or disturbances and provide good adaptability. Since the update mechanism does not depend on the model, the proposed method can be extended to nonlinear systems.

Notations

- FIE: Full information estimator
- $x(t)$: State vector
- $y$: Measurement outputs
- $\Psi(x)$: Objective function
- $G$: Noise weighting matrix
- $w$: Process noise vector
- $v$: Measurement noise vector
- $f$: Process vector
- $h$: Measurement vector
- $A$: Jacobian matrix with respect to $x$
- $C$: Jacobian matrix of $h$ with respect to $x$
- $N$: Horizon length
- $\Gamma(\cdot)$: Arrival cost
- $Q$: Weighting matrix representing the confidence in the dynamic model
- $R$: Weighting matrix representing the confidence in the measurements
- $\bar{x}$: Estimate vector
- $\bar{x}_{0:t}$: Estimate vector in the FIE problem
- $\bar{x}_0$: Initial state
- $\bar{x}_{t-N}$: Initial state
- $P_{t-N}^\ell$: Weighting matrix
- $\lambda^\ell$: Forgetting factor
- $\alpha^\ell$: Dead zone function
- $N_0$: Design parameter
- $e^\ell$: Estimation error
- $d_{1,2}$: Design parameters.

Data Availability

No data were used to support the findings of this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] A. Alessandri, M. Baglietto, G. Battistelli, and M. Gaggero, “Moving-horizon state estimation for nonlinear systems using neural networks,” IEEE Transactions on Neural Networks, vol. 22, no. 5, pp. 768–780, 2011.
[2] 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.
[3] 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.
[4] G. Battistelli, “Distributed moving-horizon estimation with arrival-cost consensus,” IEEE Transactions on Automatic Control, vol. 64, no. 8, pp. 3316–3323, 2018.
[5] R. Schneider and W. Marquardt, “Convergence and stability of a constrained partition-based moving horizon estimator,” IEEE Transactions on Automatic Control, vol. 61, no. 5, pp. 1316–1321, 2015.
[6] X. Yin and J. Liu, “Event-triggered state estimation of linear systems using moving horizon estimation,” IEEE Transactions on Control Systems Technology, vol. 29, no. 2, pp. 901–909, 2020.
[7] X. Yin and J. Liu, “Distributed moving horizon state estimation of two-time-scale nonlinear systems,” Automatica, vol. 79, pp. 152–161, 2017.
[8] 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.
[9] E. L. Haseltine and J. B. Rawlings, “Critical evaluation of extended Kalman filtering and moving-horizon estimation,” Industrial & Engineering Chemistry Research, vol. 44, no. 8, pp. 2451–2460, 2005.
[10] Q. Liu and Z. Wang, “Moving-horizon estimation for linear dynamic networks with binary encoding schemes,” IEEE Transactions on Automatic Control, vol. 66, no. 4, pp. 1763–1770, 2020.
[11] 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.
[12] P. Küh, 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.
[13] C. V. Rao and J. B. Rawlings, “Constrained process monitoring: moving-horizon approach,” AIChE Journal, vol. 48, no. 1, pp. 97–109, 2002.
[14] E. Chu, A. Keshavarz, D. Gorinevsky, and S. Boyd, “Moving horizon estimation for staged qp problems,” in Proceedings of the 2012 IEEE 51st IEEE Conference on Decision and Control (CDC), pp. 3177–3182, IEEE, Maui, HI, USA, December 2012.
[15] U. Sridhar, “Computing arrival cost parameters in moving horizon estimation using sampling based filters,” Journal of Process Control, vol. 19, no. 9, pp. 1576–1588, 2009.
[16] A. A. Al-Matouq and T. L. Vincent, “Multiple window moving horizon estimation,” *Automatica*, vol. 53, pp. 264–274, 2015.

[17] L. Zou, Z. Wang, J. Hu, and D. Zhou, “Moving horizon estimation with unknown inputs under dynamic quantization effects,” *IEEE Transactions on Automatic Control*, vol. 65, no. 12, pp. 5368–5375, 2020.

[18] L. Zou, Z. Wang, H. Dong, and Q. L. Han, “Moving horizon estimation with multirate measurements and correlated noises,” *International Journal of Robust and Nonlinear Control*, vol. 30, no. 17, pp. 7429–7445, 2020.

[19] C. Song and T. E. Bullock, “Parameter estimation algorithm with a variable forgetting factor applied to adaptive control,” *IFAC Proceedings*, vol. 26, no. 2, pp. 359–362, 1993.

[20] X. Hu, H. Jiang, F. Feng, and B. Liu, “An enhanced multi-state estimation hierarchy for advanced lithium-ion battery management,” *Applied Energy*, vol. 257, Article ID 114019, 2020.

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

[22] S. C. Patwardhan, S. Narasimhan, P. Jagadeesan, B. Gopaluni, and S. L. Shah, “Nonlinear Bayesian state estimation: a review of recent developments,” *Control Engineering Practice*, vol. 20, no. 10, pp. 933–953, 2012.

[23] K. R. Muske, J. B. Rawlings, and J. H. Lee, “Receding horizon recursive state estimation,” in *Proceedings of the 1993 American Control Conference*, pp. 900–904, IEEE, San Francisco, CA, USA, June 1993.

[24] 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.

[25] J.-P. Corriou, “Parametric estimation algorithms,” in *Process Control*, pp. 455–503, Springer, Heidelberg, Germany, 2018.

[26] C. R. Johnson, “Positive definite matrices,” *The American Mathematical Monthly*, vol. 77, no. 3, pp. 259–264, 1970.

[27] T. R. Fortescue, L. S. Kershenbaum, and B. E. Ydstie, “Implementation of self-tuning regulators with variable forgetting factors,” *Automatica*, vol. 17, no. 6, pp. 831–835, 1981.

[28] M. A. Müller, “Nonlinear moving horizon estimation for systems with bounded disturbances,” in *Proceedings of the 2016 American Control Conference (ACC)*, pp. 883–888, Boston, MA, USA, July 2016.

[29] B. Boulkroune, M. Zasadzinski, and M. Darouach, “Moving horizon state estimation for linear discrete-time singular systems,” *IET Control Theory & Applications*, vol. 4, no. 3, pp. 339–350, 2010.

[30] A. E. A. Joel, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, “Casadi: a software framework for nonlinear optimization and optimal control,” *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1–36, 2019.