Distributed moving horizon estimation with pre-estimating observer

Antonello Venturino∗,†, Sylvain Bertrand †, Cristina Stoica Maniu∗, Teodoro Alamo‡ and Eduardo F. Camacho‡

Abstract—This paper proposes a new Distributed Moving Horizon Estimation (DMHE) algorithm for state estimation of a discrete-time linear system by a sensor network. The main contribution consists in using a pre-estimating Luenberger observer in the formulation of the local problem to be solved by each sensor. This results in a significant reduction of the computation time while preserving at the same time the accuracy of the estimates. The state estimate computed by each sensor is capable to converge even under weak observability conditions. Comparison with existing approaches is proposed through a simulation example, for both centralized and distributed cases, in terms of estimation accuracy and computation time.

Index Terms—distributed state estimation, moving horizon estimation, sensor networks, Luenberger observers, linear systems.

1. INTRODUCTION

In recent decades, there has been an increasing interest in research on distributed state estimation due to its variety of applications over sensor networks such as target tracking [1]–[2], exploration [3], surveillance [4], etc. The estimation algorithms used in these applications are mainly formulated as centralized fusion architectures, where all the sensors transmit their local measurements to a central unit that processes the provided measurements to update the estimate. In general, the centralized algorithms are not scalable, since with the increasing number of sensors the complexity of the problem to solve increases, too. Furthermore, in case of large-scale networks, communication with a unique central unit can hardly be achieved e.g. due to communication bandwidth limitations. Unlike centralized schemes, in the distributed approaches [5]–[7] each sensor computes a local estimate using information acquired by other connected sensors within a local neighborhood. This can improve robustness to sensor failure exploiting redundancy and also lower the communication burden since data is transmitted amongst local nodes in the network.

The continuous decreasing costs of sensors is making these applications realizable, even tough there are still open problems to face with. In fact, the distributed algorithms need to have particular properties in order to make them attractive for the industrial community. Thus, in the context of large-scale systems, the algorithms must be scalable to be able to deal with large networks, must have low computation load since low-cost sensors are not powerful devices, must minimize the utilization of communication resources, etc. In [8], the authors reviewed several works on distributed state estimation over low-cost sensors networks, pointing out their characteristics, advantages, and challenging issues.

In recent years, Moving Horizon Estimation (MHE) techniques and the distributed counterpart (DMHE) have been used to successfully deal with large sensor networks [9]. The first idea of MHE proposed in [10] consists in estimating the current states by solving a least-square optimization problem penalizing on one hand the deviation between acquired and predicted measurements, and on the other hand the distance between the estimated state and some a priori information on it. MHE is a practical strategy for constrained state estimation and a lot of research has been devoted to develop stability guarantees on the estimation error, e.g., [11]–[13]. Although this approach is functional for control engineers offering the freedom to tune the parameters of the cost function, a strength and a weakness of this approach is the use of an optimization problem. This has to be solved within the sampling time; however, for large-scale systems this issue becomes critical. There have been several attempts in trying to reduce the computation demand of MHE. One idea is to add a pre-estimating observer in the formulation. The authors of [12] proposed a MHE strategy with a Luenberger observer that leads to good performance especially for large estimation horizons. They also propose an optimization problem to tune the parameters minimizing the effects of measurement noise and model errors. In [14], the formulation is generalized using a weight matrix for the penalty function and adding states constraints. A MHE with pre-estimation has been proposed also for non-linear systems [15]. In [6] a distributed algorithm is proposed for linear systems with a stability proof under weak observability conditions (exploiting a consensus-on-estimate and a consensus weight term in the DMHE formulation). However, the computation time issue becomes crucial, since usually the network could be composed of low-cost sensors.

In this paper we propose a DMHE algorithm with pre-estimation, based on the ideas of [6], [12], [14]. A pre-estimating Luenberger observer is considered in the formulation of the local problem to be solved by each sensor, resulting in a significant reduction of the computation time. The main contribution of this paper covers: i) reducing the computation time of the optimization problems, ii) preserving the accuracy of the estimation errors. This allows the use of this type of algorithms for time-sensitive applications.
The paper is organized as follows. Section II introduces
the main theoretical statements and definitions. Sections III
and IV present the proposed versions of centralized MHE
and distributed MHE approaches, respectively. In Section V
a simulation example is proposed, while Section VI provides
some concluding remarks.

II. NOTATIONS AND DEFINITIONS

Consider the following discrete-time linear time invariant
(LTI) system:

$$x_{t+1} = Ax_t + w_t,$$  \hspace{1cm} (1)

where $x_t \in \mathbb{X} \subseteq \mathbb{R}^n$ is the state vector and $w_t \in \mathcal{W} \subseteq \mathbb{R}^n$
is a white noise with covariance matrix $Q$. The sets $\mathbb{X}$ and $\mathcal{W}$
are assumed convex with $0 \in \mathbb{X}$ and $0 \in \mathcal{W}$. The initial
state $x_0$ is assumed and assumed to be modelled by a random
variable. Let us denote respectively by $\mu$ and $\Pi_0$ its mean and
covariance matrix. The measurements are carried out by $M$
sensors, that could be different from each other:

$$y^j_t = C^j x_t + v^j_t, \hspace{1cm} i = 1, \ldots, M,$$  \hspace{1cm} (2)

where $v^j_t \in \mathbb{R}^p_i$ is a white noise with covariance $R^j$.

Consider a network described by the directed graph
$\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where nodes in $\mathcal{N} = \{1, 2, \ldots, M\}$
represent the sensors and the edge $(j, i) \in \mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$
represents the communication link from sensor $j$ to sensor $i$. We assume
that all the links have a self-loop $(i, i) \in \mathcal{E}, \forall i \in \mathcal{N}$. The
transmission neighbourhood $N^j_i$ of the sensor $i$ is $N^j_i = \{j \in \mathcal{N} : (i, j) \in \mathcal{E}\}$, the reception neighbourhood $N^j_0$ of
the sensor $i$ is $N^j_0 = \{j \in \mathcal{N} : (j, i) \in \mathcal{E}\}$ and the regional
neighbourhood $N^i$ of the sensor $i$ is $N^i = N^j_i \cup N^j_0$. The
number of nodes $j \in N^j_0$ is denoted by $M^j$.

A stochastic matrix $K \in \mathbb{R}^{M \times M}$ is associated to the graph
$\mathcal{G}$ such that the elements:

$$k_{ij} > 0 \hspace{1cm} \text{if} \hspace{1cm} (j, i) \in \mathcal{E}, \hspace{1cm} (3a)$$

$$k_{ij} = 0 \hspace{1cm} \text{otherwise}, \hspace{1cm} (3b)$$

$$\sum_{j=1}^{M} k_{ij} = \mathbb{1}, \hspace{1cm} \forall i = 1, \ldots, M. \hspace{1cm} (3c)$$

The topology of the sensor network is assumed to be time-
invariant thus the matrix $K$ is constant. This matrix will be
used to compute the consensus terms in the DMHE algorithm
described in Section IV.

We use the notation introduced in [6] to recognize local, regional
and collective information. Considering Sensor $i$, an
information is considered to be local if it is related only to
the node $i$, regional if it is related to the nodes in $N^j_i$, and
collective if it is related to the entire network. Corresponding
notations from [6] are used to distinguish local, regional and
collective variables. For a given variable $z, z'$ denotes its local
version, $z$ the regional one and $Z$ the collective one. Using
these notations, the regional measurements of a sensor $i$ with
reception neighborhood $N^j_i = \{j_1, \ldots, j_{M^j_i}\}$ are

$$\hat{y}^j_i = C^j x_t + \hat{v}^j_i,$$  \hspace{1cm} (4)

where $\hat{y}^j_i = [(y_1^j)^T, \ldots, (y_{M^j_i}^j)^T]^T \in \mathbb{R}^{p_i}$ with $\hat{v}_i = \sum_{i \in N^j_i} p_i \hat{v}_i, \hspace{1cm} C^j_i = [(C^{j_1})^T, \ldots, (C^{j_{M^j_i}})^T]^T$ and $\hat{v}^j_i =

$$[(\hat{v}_{i1}^{j_1})^T, \ldots, (\hat{v}_{iM^j_i}^{j_{M^j_i}})^T]^T.$$

The covariance matrix associated to the region $\hat{v}^j_i$ of sensor $i$ is $R^i = diag(R^i_1, \ldots, R^i_{M^j_i}).$

The system is locally observable by sensor $i$ (sensor $i$ is locally observable) if the pair $(A, C^i)$ is observable.
The system is regionally observable by sensor $i$ (sensor $i$ is
regionally observable) if the pair $(A, C^i)$ is observable.
The system is collective observable if the pair $(A, C)$ is observable.

III. CENTRALIZED MHE WITH PRE-ESTIMATION

Moving Horizon Estimation with pre-estimation is based on
the idea in [12] and [14], where the authors proposed a MHE
formulation relying on a Luemberger observer. The centralized
scheme makes use of the collective information solely.

In this strategy, for a given horizon length $N \geq 1$, an estimate $\hat{x}_{t|t}$ of the state at time $t$ is computed by solving
the constrained minimization problem, hereafter denoted by
MHE pre:

$$\hat{x}_{t-N|t} = \arg \min_{x_{t-N}} J(t - N, t, \hat{x}_{t-N}, \tilde{\Psi}, \Gamma_{t-N}) \hspace{1cm} (5)$$

$$\text{s.t.} \hspace{1cm} \dot{x}_{k+1} = Ax_k + L\tilde{v}_k, \hspace{1cm} (6)$$

$$\tilde{y}_k = Cx_k + \hat{v}_k, \hspace{1cm} (7)$$

$$\hat{x}_{k} \in \mathbb{X}, \hspace{1cm} (8)$$

$$\forall \hspace{0.1cm} k = t - N, \ldots, t. \hspace{1cm} (9)$$

The gain matrix $L \in \mathbb{R}^{n \times p}$ in (6), with $p = \sum_{i=1}^{M} p_i$, is chosen such that $\Phi = A - LC$ is Schur stable. The cost function $J$ given by:

$$J(t - N, t, \hat{x}_{t-N}, \tilde{\Psi}, \Gamma_{t-N}) = \frac{1}{2} \sum_{k=t-N}^{t} \| \tilde{v}_k \|^2_{R^{-1}} + \Gamma_{t-N} (\hat{x}_{t-N}, \hat{x}_{t-N|t-1}),$$  \hspace{1cm} (9)

where $R = diag(R^1, \ldots, R^M)$. The so-called initial penalty
function $\Gamma_{t-N}(\hat{x}_{t-N}, \hat{x}_{t-N|t-1})$ in (9) is defined as follows:

$$\Gamma_{t-N}(\hat{x}_{t-N}, \hat{x}_{t-N|t-1}) = \frac{1}{2} \| \hat{x}_{t-N} - \hat{x}_{t-N|t-1} \|^2_{\Pi_{t-N|t-1}}^{-1}, \hspace{1cm} (10)$$

where $\hat{x}_{t-N|t-1}$ is the second estimated state of the sequence
computed at the previous time. The positive-definite symmetric
weighting matrix $\Pi_{t-N|t-1}$ in (10) is the unique solution of
the following discrete-time algebraic Riccati equation [15]:

$$\Pi_{t-N|t-1} = Q + A\Pi_{t-N-1|t-2}A^T -$$

$$- A\Pi_{t-N-1|t-2}C^T (R + C\Pi_{t-N-1|t-2}C^T)^{-1} C\Pi_{t-N-1|t-2}A^T,$$  \hspace{1cm} (11)

subject to the initial condition $\Pi_0$. When $t \leq N$, one sets $N = t$. We denote by $\hat{x}_{k|t}^{i} | k = t - k$ the sequence of the estimated states computed using (6) with the optimum solution $\hat{x}_{t-N|t}$ of (5) as initial condition. The last term $\hat{x}_{k+1}$ of this sequence corresponds to the estimate of the state $x_t$ of the system at
the current time instant $t$. 

175
IV. DISTRIBUTED MHE WITH PRE-ESTIMATION

In [6] the authors proposed a distributed Moving Horizon Estimation method to estimate the state and the input of the model (1). In their strategy, the optimization problem to be solved online at time $t$ involves the computation of the state trajectory over the past horizon. As in classical MHE schemes, this computation is done by propagating forward the state, from its initial condition at $t - N$, using the dynamic model of the system. This can accumulate the estimation error, especially when $N$ is large and the system is unstable. In this section, we propose a new DMHE strategy with pre-estimation by introducing a Luenberger observer in its formulation. This will mitigate the effect of model uncertainty in the a priori estimate and thus will contribute to enhance the estimation accuracy. This improvement will also reduce the computation time required to solve the optimization problem, because the number of optimization variables involved is lower and does not depend on the horizon length $N$. Further details are given in the next sections.

A. Local minimization problem

Here we formulate the proposed DMHE scheme, from now on denoted as DMHE$_{pre}$, where each sensor $i \in N$ solves its own local moving horizon estimation problem based on regional measurements $\hat{y}_k^i$ and some shared information amongst the neighbourhood $N_i$. For a given horizon length $N \geq 1$, each sensor $i \in N$ computes an estimate $\hat{x}_{t-N}^i$ of the state $x_t$ at time $t$ by solving the following constrained minimization problem with pre-estimation, hereafter denoted as DMHE$_{t-pre}$:

$$\hat{x}_{t-N}^i = \arg \min_{\hat{x}_{t-N}^i \in \mathcal{X}} J^i(t-N,t,\hat{x}_{t-N}^i, \hat{\theta}^i, \Gamma_{t-N}^i)$$

subject to

$$\hat{x}_{t+1}^i = A\hat{x}_t^i + C\hat{\theta}_t^i,$$

$$\hat{y}_k^i = C\hat{x}_k^i + \hat{\nu}_k^i,$$

$$\forall k = t-N, \ldots, t.$$ (12)

Note that the equations remind the ones of the MHE$_{pre}$ problem but using regional information instead. There are also some differences due to the consensus terms embodied in the initial penalty, which will be discussed in detail later on. First of all, the Luenberger gain $L^i$ is computed such that $\Phi^i = A - L^i C^i$ is Schur stable when the sensor $i$ is regionally observable, i.e. the pair $(A, C^i)$ is observable. Otherwise, as extrema ratio, one can compute $L^i$ in order to minimize the propagation of the error along the prediction horizon by keeping the spectrum radius of $\Phi^i$ as small as possible.

The cost function $J^i$ is given by:

$$J^i(t-N, t, \hat{x}_{t-N}^i, \hat{\theta}^i, \Gamma_{t-N}^i) = \frac{1}{2} \sum_{k=t-N}^{t} \left\| \hat{\nu}_k^i \right\|^2_{(\Gamma_{t-N}^i)^{-1}}$$

$$+ \Gamma_{t-N}^i (\hat{x}_{t-N}^i - \hat{x}_{t-N}|t-1),$$

where the initial penalty function $\Gamma_{t-N}^i(\hat{x}_{t-N}^i, \hat{x}_{t-N}|t-1)$ in (16) is defined as follows:

$$\Gamma_{t-N}^i(\cdot) = \frac{1}{2} \left\| \hat{x}_{t-N}^i - \hat{x}_{t-N}|t-1 \right\|^2_{(\Pi_{t-N}|t-2)^{-1}}.$$ (17)

Let $\hat{x}_{t-N}|t-1$ denote the weighted average of the state estimates produced by sensors $j \in N_j$, computed as follows:

$$\hat{x}_{t-N}|t-1 = \sum_{j \in N_j} k_{ij} \hat{x}_{t-N}|t-1,$$ (18)

where $\hat{x}_{t-N}|t-1$ is the second estimated state in the sequence computed at the previous time by sensor $j$. A consensus-like term on the estimates is therefore introduced by (18) in the initial penalty (17). It helps to improve the accuracy of the local estimates and enables to guarantee convergence of the estimates to the state of the system even without regional observability [6].

The term $\Pi_{t-N}|t-1$ is computed as in [6]. For the sake of clarity, we recall here the procedure to compute it by:

$$\Pi_{t-N}|t-1 = \sum_{j \in N_j} M^j k_{ij} \Pi_{t-N}|t-1,$$ (19)

where the update of $\Pi_{t-N}|t-1$ is performed by the sensor $i$ based on regional information. More specifically, the matrix $\Pi_{t-N}|t-1$, with $i \in N$, is given by one iteration of the difference Riccati equation associated to a Kalman filter for the system:

$$\left\{ \begin{array}{l}
\dot{x}_t = Ax - t - N + w_t - N - 1 \\
\hat{z}_{t-N}^i = \tilde{O}_N^i x_t + V_{t-N}^i
\end{array} \right.$$ (20)

where $V_{t-N}^i$ represents the measurements noise and $\tilde{O}_N^i$ defines the $i$-th sensor regional observability matrix:

$$\tilde{O}_N^i = [\begin{array}{ccc}
(C^i)^T & (C^i A)^T & \cdots & (C^i A^{N-1})^T\end{array}]^T$$

If we define

$$C_N^i = \begin{bmatrix}
0 & 0 & \cdots & 0 \\
C^i & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
C^i A^{N-2} & C^i A^{N-3} & \cdots & C^i
\end{bmatrix} \in \mathbb{R}^{p_i \times n(N-1)},$$ (21)

$$R_N^i = \text{diag}(R_1^i, \ldots, R_i^i) \in \mathbb{R}^{p_i \times p_i},$$ (22)

$$Q_{N-1} = \text{diag}(Q_1, \ldots, Q_i) \in \mathbb{R}^{n(N-1) \times n(N-1)},$$ (23)

$$\text{Cov}[V_{t}^i] = R_N^i + C_N^i Q_{N-1} C_N^i T,$$ (24)

and set the covariance of the estimate $\hat{x}_{t-N}^i$ as

$$\Pi_{t-N-1}^i = \left( \Pi_{t-N-1}^i \right)^{-1} + (C_N^i)^T (R_N^i)^{-1} C_N^i,$$ (25)
the resulting Riccati recursive equation is given by
\[
\Pi_{i}^{t-N|t-1} = R^{i}(\Pi_{i}^{t-N-1|t-2}, Q, \bar{R}_{i}^{N})
\]
\[
= A\Pi_{i}^{t-N-1|t-2}A^{T} + \cdots
\]
pointed out in [6], and recalled next, this example is challenging due to the non regional observability of some of

Since the communication network topology is assumed to be time-invariant, these equations can be computed off-line.

B. Network information exchange

It is worth highlighting the way each node exchanges the information within its neighbourhood. For this purpose, we recall here some assumptions that play a major role:

- the sensors can be different from each other;
- the sensors characteristics (noise covariance and type of measurements) are not time-varying;
- the network topology is not time-varying;
- each sensor \( i \) knows its reception neighbourhood \( N_{i}^{\rho} \);
- no delay in the communication network.

This implies that the matrices \( C_{i}^{\rho} \) can differ for all \( i \in N \). Since the reception neighbourhood \( N_{j}^{\rho} \) is known \textit{a priori}, it is not necessary to exchange the information on the matrices \( C_{i}^{\rho} \) and \( R^{i} \) at each time. Moreover, this allows one to compute off-line the Luenberger gains \( L^{i} \).

C. DMHE\textsubscript{pre} procedure

Finally, we have all the elements to describe the procedure of the proposed distributed scheme.

1) Off-line computation:

a) Since the Luenberger gains \( L^{i} \) are constants, they can be computed off-line.

b) All nodes store the estimate \( \hat{x}_{0|0} = \hat{x}_{0} = \mu \) of \( x_{0} \), where \( \mu \) is given, and the covariance matrix \( \Pi_{0} \) of \( x_{0} \).

2) Initialization:

a) At the first step \( t = 0 \), each node \( i \in N \) performs a first measurement \( y_{0}^{i} \) and sends the information to the nodes \( j \in \mathcal{N}_{i}^{0} \).

3) Online computation: each node \( i \in N \)

a) performs the measurement \( y_{t}^{i} \) and sends it to the nodes \( j \in \mathcal{N}_{i}^{\rho} \);

b) if \( 1 \leq t \leq N \), then \( N = t \), \( \hat{\Pi}_{i}^{t-N|t-1} = \hat{\Pi}_{i}^{0|t-1} = \Pi_{0} \) and \( \hat{x}_{0|t-1}^{i} = \hat{x}_{0|t-1}^{0} \);

c) if \( t > N \):
   i) computes \( \hat{\Pi}_{i}^{t-N|t-1} \) according to (24), (25) and (26), and sends it to the nodes \( j \in \mathcal{N}_{i}^{\rho} \);
   ii) computes \( \hat{\Pi}_{i}^{t-N|t-1} \) according to (19);
   iii) solves the problem DMHE\textsubscript{pre} and store the solution \( \hat{x}_{t|t}^{i} \);
   e) sends \( \hat{x}_{t-N+1|t}^{i} \) to the nodes \( j \in \mathcal{N}_{i}^{\rho} \).

Notice that the steps 3(c) and 3(c)ii in the DMHE\textsubscript{pre} procedure could be computed off-line. However, for clarity reasons with respect to the exchanging information timing, they have been described in the online step.

V. Example

A performance analysis is provided in this section for the proposed DMHE algorithm with pre-estimation through a simulation example. Comparison is performed to the centralized MHE of [11] and the DMHE of [6]. The system considered for illustration is the one used in [6] an given by

\[
x_{t+1} = \begin{bmatrix}
0.9962 & 0.1949 & 0 & 0 \\
-0.1949 & 0.3819 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & -1.21 & 1.98
\end{bmatrix} x_{t} + w_{t},
\]

with \( x_{t} = [x_{1,t}, x_{2,t}, x_{3,t}, x_{4,t}]^{T} \) is the state vector and \( w_{t} \in \mathbb{R}^{4} \) is a white noise with covariance \( Q = \text{diag}(0.0012, 0.038, 0.0012, 0.038) \). Notice that the system is unstable since the eigenvalues of \( A \) are 0.9264, 0.4517, 0.99 ± 0.4795i and [0.99 ± 0.4795i] > 1. For this example, one considers \( \mathcal{X} = \mathbb{R}^{4} \).

The initial values of the algorithms have been set as \( \mu = [0 \ 0 \ 0 \ 0]^{T} \), \( \Pi_{0} = I_{4} \) and \( N = 5 \).

To compare the results of all algorithms we use \( M = 4 \) sensors both for the centralized and distributed cases. For the centralized scheme the following measurement equation is considered:

\[
y_{t} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0
\end{bmatrix} x_{t} + v_{t},
\]

with \( \text{Var}(v_{t}) = R = I_{4} \).

For the distributed schemes, the following measurement equations are used:

\[
y_{t}^{i} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix} x_{t} + v_{t}^{i} \quad \text{if } i = 1, 2,
\]

\[
y_{t}^{i} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix} x_{t} + v_{t}^{i} \quad \text{if } i = 3, 4,
\]

where \( \text{Var}(v_{t}^{i}) = R^{i} = 1, \ i = 1, \ldots, 4 \). The nodes are connected as reported by the graph in Fig. 1 and the associated matrix is defined as follows:

\[
K = \begin{bmatrix}
0.5 & 0 & 0 & 0.5 \\
0.5 & 0.5 & 0 & 0 \\
0 & 0.5 & 0 & 0.5 \\
0 & 0 & 0.5 & 0.5
\end{bmatrix}
\]

Fig. 1. Communication network.

As pointed out in [6], and recalled next, this example is challenging due to the non regional observability of some of
the sensors. In fact, analyzing the sensor 1, we notice that at each time instant it has available the information about \(x_{1,t}\), directly measured, and \(x_{3,t}\), transmitted by sensor 4. Likewise, the information available to sensor 2 consists of \(x_{1,t}\), directly measured, and \(x_{3,t}\), transmitted by sensor 1. The same applies to the sensors 3 and 4, which have information about \((x_{1,t}, x_{3,t})\) and \((x_{3,t}, x_{3,t})\), respectively. Consequently, the sensors 1 and 3 are regionally observable, while the sensors 2 and 4 are not, since the pairs \((A, C^2)\) and \((A, C^4)\) are not observable.

The Luenberger gains, both for MHE\(_{\text{pre}}\) and DMHE\(_{\text{pre}}\), have been chosen such that the eigenvalues of \(\Phi = A - LC\) for the centralized scheme, and \(\Phi^t = A - L'C^t\) for the distributed one, are equal to the values shown in the Table I. These values have been chosen to ensure that matrices \(\Phi\), \(\Phi^1\) and \(\Phi^3\) are Schur stable, and to keep the spectrum radius of \(\Phi^2\) and \(\Phi^4\) "small" enough.

<table>
<thead>
<tr>
<th></th>
<th>(\lambda_1)</th>
<th>(\lambda_2)</th>
<th>(\lambda_3)</th>
<th>(\lambda_4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHE(_{\text{pre}})</td>
<td>0.9</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td>DMHE(_{1-\text{pre}})</td>
<td>0.9</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td>DMHE(_{2-\text{pre}})</td>
<td>0.45</td>
<td>0.58</td>
<td>0.90 + 0.48i</td>
<td>0.99 - 0.48i</td>
</tr>
<tr>
<td>DMHE(_{3-\text{pre}})</td>
<td>0.9</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td>DMHE(_{4-\text{pre}})</td>
<td>0.93</td>
<td>0.45</td>
<td>1.17 + 0.98i</td>
<td>1.17 - 0.98i</td>
</tr>
</tbody>
</table>

We consider two different performance metrics for the evaluation of the algorithms: computation time \(\tau_c\), of which we examine sum, minimum and maximum time, and root mean square error (RMSE):

\[
\text{RMSE} = \left( \frac{1}{t_f - t_c} \sum_{t=t_c}^{t_f} \| e_t \|^2 \right)^{1/2}, \quad (28)
\]

where \(t_c\) is the convergence of the algorithms, \(t_f\) is the final time instant, with \(t_f > t_c\), and \(e_t = x_t - \hat{x}_t\) is the estimation error. In this example, the sampling period is chosen to be 1s, while \(t_f = 20s\) corresponds to the simulation duration.

The simulation has been carried out by using a setup implemented within the MATLAB R2019b environment and the solver LMI Lab [16] over a Linux Ubuntu 20.04 PC equipped with an Intel Core i7-7700HQ, 2.80 GHz.

Fig. 2 shows the evolution of the system (27) and the estimates by MHE\(_{\text{pre}}\) and DMHE\(_{\text{pre}}\) algorithms. It is worth noting that the last two states have unstable dynamics and are larger in terms of magnitude with respect to the first two, nevertheless one can consider that the convergence time is \(t = 5s\).

Estimation errors corresponding to MHE\(_{\text{pre}}\) and DMHE\(_{\text{pre}}\) are represented in Fig. 3. It can be notice that the MHE\(_{\text{pre}}\) compensates the initial estimation error very fast, within a few iterations, since the system is collective observable. As pointed out in [6] for the DMHE, when using the DMHE\(_{\text{pre}}\) the estimates produced by sensor 2 (respectively 4) relative to the states \(x_{3,t}, x_{4,t}\) (respectively \(x_{1,t}, x_{2,t}\)) also exhibit large errors for \(t < 5\). In fact, as previously mentioned, these states cannot be observed by these sensors using regional information. Effectiveness of the consensus terms also introduced in the DMHE\(_{\text{pre}}\) scheme is therefore illustrated here since all the estimation errors tend to converge to the same value despite the weak observability conditions of this example.

For the sake of clarity, in Fig. 2 and Fig. 3 the estimates and the estimation errors, respectively, from MHE and DMHE are omitted. It is also because \(e_t\) and \(\hat{x}_t\) from MHE and DMHE are very similar with the ones from MHE\(_{\text{pre}}\) and DMHE\(_{\text{pre}}\), respectively, as can be checked by comparing Fig. 3 with Fig. 4(a) in [6, p. 2468].
Table II shows the performance metrics that have been taken into account. In particular, the computation time is always lower for our proposed algorithms with respect to the ones without pre-estimation. This not come at a cost in terms of accuracy, because the RMSEs produced are similar amongst them. The maximum computation time for the algorithms without pre-estimation is greater than the sampling period. This means that there are some time steps where real time feasibility is not obtained. This can be checked on Fig. 4. Pre-estimation enables to reduce computation time by a factor close to 5 and would enable real time implementation without requiring fast optimization or ad hoc implementation.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>( \tau )</th>
<th>( \tau )</th>
<th>( \tau )</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHE ( \text{pre} )</td>
<td>0.27</td>
<td>1.13</td>
<td>17.25</td>
<td>1.4</td>
</tr>
<tr>
<td>DMHE ( 1 )</td>
<td>0.24</td>
<td>0.27</td>
<td>5.21</td>
<td>0.99</td>
</tr>
<tr>
<td>DMHE ( 1 ) ( \text{pre} )</td>
<td>0.26</td>
<td>1.10</td>
<td>17.30</td>
<td>0.74</td>
</tr>
<tr>
<td>DMHE ( 2 ) ( \text{pre} )</td>
<td>0.24</td>
<td>0.29</td>
<td>5.20</td>
<td>0.84</td>
</tr>
<tr>
<td>DMHE ( 2 )</td>
<td>0.26</td>
<td>1.02</td>
<td>16.30</td>
<td>1.53</td>
</tr>
<tr>
<td>DMHE ( 3 ) ( \text{pre} )</td>
<td>0.24</td>
<td>0.27</td>
<td>5.09</td>
<td>1.63</td>
</tr>
<tr>
<td>DMHE ( 3 )</td>
<td>0.27</td>
<td>1.08</td>
<td>16.49</td>
<td>0.77</td>
</tr>
<tr>
<td>DMHE ( 4 ) ( \text{pre} )</td>
<td>0.25</td>
<td>0.29</td>
<td>5.18</td>
<td>0.76</td>
</tr>
<tr>
<td>DMHE ( 4 )</td>
<td>0.26</td>
<td>1.10</td>
<td>17.24</td>
<td>0.90</td>
</tr>
<tr>
<td>DMHE ( 4 ) ( \text{pre} )</td>
<td>0.24</td>
<td>0.27</td>
<td>5.11</td>
<td>1.08</td>
</tr>
</tbody>
</table>

Figure 4 shows the computation time \( \tau \) of all algorithms. In this figure is more clear that our proposed schemes are less time demanding. In fact, the computation time mainly relies on the number of optimization parameters \( n_{\text{op}} \), the evaluation of the cost and constraints. As it can be seen, \( n_{\text{op}} \) is the most important, since the algorithms without pre-estimation have \( n_{\text{op}} = n(N + 1) \), while the ones with pre-estimation have only \( n_{\text{op}} = n \). Thus, the latter ones do not depend on the horizon length \( N \). Since \( N = t \) for \( t \leq 5 \) and \( N = 5 \) for \( t > 5 \), as excepted, the computation time rises as \( N \) increases for MHE and DMHE.

VI. CONCLUSION

We presented a novel algorithm based on Moving Horizon Estimation (MHE) concept for distributed state estimation of discrete-time linear time-invariant systems. The use of a pre-estimation observer results in a significant reduction of the computation time. The proposed strategy has been validated via an illustrative numerical example. The estimation errors produced by the DMHE \( \text{pre} \) are capable to converge even if some sensors in the network are not observable, due to the consensus terms embodied in the optimization problem. Moreover, the accuracy of the estimation errors is preserved (together with a reduced computation time), in the sense that it is comparable with the one of the original formulation [6]. Further improvements will include stability analysis of the estimation error provided by the proposed DMHE strategy with pre-estimation.

ACKNOWLEDGMENT

The authors acknowledge support of MEyC Spain (contract PID2019-106212-RB-C41), the European Research Council and the UE ERDF (Advanced Grant OCONTOSOLAR, Project ID: 789051).

REFERENCES