2010 American Control Conference
Marriott Waterfront, Baltimore, MD, USA
June 30-July 02, 2010

State estimation for large-scale partitioned systems: a moving horizon approach

Marcello Farina, Giancarlo Ferrari-Trecate, Riccardo Scattolini

Abstract— **In this paper we propose novel state-estimation methods for large-scale discrete-time constrained linear systems that are partitioned, *i.e.* made by coupled subsystems with non-overlapping states. We focus on moving horizon estimation (MHE) schemes due to their capability of exploiting physical constraints on states and noise in the estimation process. We propose three different partition-based MHE (PMHE) algorithms where each subsystem solves reduced-order MHE problems to estimate its own state. Different estimators have different computational complexity, accuracy and transmission requirements among subsystems. Numerical simulations demonstrate the viability of our approach.**

I. Introduction

State-estimation problems for large-scale systems decomposed into physically coupled subsystems has received great attention within the control community since the 70's. The main motivation is that decentralized state-estimation enables the development of output-feedback decentralized control schemes that are ubiquitous in numerous applications such as power systems [20], transport networks [16] and process control [19].

Many studies focused on the design of decentralized Kalman filters. Early works, *e.g.* [9], [12] aimed at reducing the computational complexity of centralized Kalman filtering by parallelizing computations, requiring all-to-all communication and assuming each subsystem has full knowledge of the whole dynamics. In [11] the focus was on the use of reduced-order and decoupled models for each subsystem. This paper, beside neglecting coupling, exploits communication networks that are almost fully connected. Subsystems with overlapping states have been considered in [10], [18], [17], [19]. While the estimation schemes in [19] require all-to-all communication, in [10], [18], [17] the topology of the network is defined by dependencies among the states of subsystems resulting in a fully decentralized scheme.

One drawback of Kalman filtering is that known physical constraints on noise and state variables are not exploited in the estimation process. This can lead to suboptimal estimates or instability of the error dynamics [15]. In order to overcome these issues, moving horizon estimation (MHE) has been proposed for discrete-time linear [1], [13], nonlinear [2], [3], [14] hybrid systems [7], and is employed in distributed estimation schemes for linear systems [5], [4]. MHE amounts to solve at each time instant an optimization problem whose complexity scales with the number of states, inputs and the estimation horizon. These demanding computational requirements often hamper the applicability of centralized MHE schemes to large-scale systems.

In this paper we propose decentralized MHE algorithms with the goal of reducing the computational complexity of the centralized solution. More specifically, we consider linear constrained systems that are *partitioned*, *i.e.* decomposed into interconnected subsystems with no overlapping states, and propose three partition-based MHE (PMHE) schemes, named PMHE1, PMHE2 and PMHE3. In all cases each subsystem solves a reduced-order MHE problem in order to estimate its own states and transmits information to the neighboring subsystems. For all the proposed schemes we provide sufficient conditions for convergence to zero of the estimation errors. The three solutions have different features in terms of communication requirements among subsystems, accuracy and computational complexity. While PMHE1 and PMHE2 provide a decentralization of the MHE scheme proposed in [13], PMHE3 is inspired to the MHE strategy for unconstrained systems described in [1]. Moreover, compared to PMHE2, PMHE3 has lower computational complexity at the price of a loss in noise filtering performance. Interestingly, when the whole system is viewed as a single subsystem, PMHE3 generalizes the MHE scheme proposed in [1] to the case of constrained estimation.

The paper is structured as follows. Section II introduces the partitioned systems considered in the following. Section III describes the proposed MHE procedures, while convergence results are provided in Section IV. An illustrative example using a compartmental system is given in Section V. The proofs of the theorems and some generalizations can be found in [6].

II. Partitioned systems

Consider an observed process which obeys to the linear constrained dynamics

$$\mathbf{x}_{t+1} = \mathbf{A}\,\mathbf{x}_t + \mathbf{w}_t, \tag{1}$$

where $\mathbf{x}_t \in \mathbb{X} \subseteq \mathbb{R}^n$ $(0 \in \mathbb{X})$ is the state vector and the term $\mathbf{w}_t \in \mathbb{W} \subseteq \mathbb{R}^n$ $(0 \in \mathbb{W})$ represents a white noise with variance equal to $\mathbf{Q} \in \mathbb{R}^{n\times n}$. Let the sets $\mathbb{X}$ and $\mathbb{W}$ be convex. When $\mathbb{X} = \mathbb{R}^n$ and $\mathbb{W} = \mathbb{R}^n$ we say that the system is unconstrained. The initial condition $\mathbf{x}_0 \in \mathbb{R}^n$ is a random variable with mean $\mathbf{m}_{x_0}$ and covariance matrix $\mathbf{\Pi}_0$. Measurements on the state vector are performed according to the sensing model

$$\mathbf{y}_t = \mathbf{C}\,\mathbf{x}_t + \mathbf{v}_t \tag{2}$$

This research has been supported by the European 7th framework STREP project "Hierarchical and distributed model predictive control (HD-MPC)", contract number INFSO-ICT-223854.

M. Farina and R. Scattolini are with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, Via Ponzio 34/5, 20133, Milan, Italy {farina,riccardo.scattolini}@elet.polimi.it

G. Ferrari-Trecate is with the Dipartimento di Informatica e Sistemistica, Universita' degli Studi di Pavia, Via Ferrata 1, 27100 Pavia, Italy giancarlo.ferrari@unipv.it

where the term $\mathbf{v}_t \in \mathbb{R}^p$ represents white noise with variance equal to $\mathbf{R} \in \mathbb{R}^{p \times p}$.

Let the system (1) be partitioned in M low order interconnected non overlapping sub-models, where a generic sub-model has $x_t^{[i]} \in \mathbb{R}^{n_i}$ as state vector, *i.e.*, $\mathbf{x}_t = [(x_t^{[1]})^T \ldots (x_t^{[M]})^T]^T$ and $\sum_{i=1}^M n_i = n$. According to such decomposition, the state transition matrices $A^{[1]} \in \mathbb{R}^{n_1 \times n_1}, \ldots, A^{[M]} \in \mathbb{R}^{n_M \times n_M}$ of the M subsystems are diagonal blocks of $\mathbf{A}$, whereas the off-diagonal blocks of $\mathbf{A}$ define the coupling terms between subsystems, which can be seen as inputs to the individual subsystems. It results that the i-th subprocess obeys to the linear dynamics

$$x_{t+1}^{[i]} = A^{[i]} x_t^{[i]} + u_t^{[i],x} + w_t^{[i]}, \tag{3}$$

where $x_t^{[i]} \in \mathbb{X}_i \subseteq \mathbb{R}^{n_i}$ is the state vector, $u_t^{[i],x}$ collects state variables of other subsystems (and will be specified later on), and the term $w_t^{[i]} \in \mathbb{W}_i \subseteq \mathbb{R}^{n_i}$ represents a disturbance with variance equal to $Q^{[i]} \in \mathbb{R}^{n_i \times n_i}$. The sets $\mathbb{X}_i$ and $\mathbb{W}_i$ are convex by convexity of $\mathbb{X}$ and $\mathbb{W}$. The initial condition $x_0^{[i]} \in \mathbb{R}^{n_i}$ is a random variable with mean $m_{x_0}^{[i]}$ and covariance matrix $\Pi_0^{[i]}$. As far as the outputs of the subsystems are concerned, they can be assigned according to (2) and to the state partition. In this way the measurements on $x_t^{[i]}$ are performed according to the sensing model

$$y_t^{[i]} = C^{[i]} x_t^{[i]} + u_t^{[i],y} + v_t^{[i]} \tag{4}$$

where $u_t^{[i],y}$ collects the effect of the state variables of other subsystems (it will be specified later on), and the term $v_t^{[i]} \in \mathbb{R}^{p_i}$ represents white noise with variance equal to $R^{[i]} \in \mathbb{R}^{p_i \times p_i}$. Notice that, in general, some outputs of the system (1) can be considered as outputs of more than one of the process subsystems, *i.e.*, $\bar{p} = \sum_{i=1}^m p_i \geq p$. We now define $\bar{\mathbf{y}}_t = [(y_t^{[1]})^T \ldots (y_t^{[M]})^T]^T$. Accordingly, there exists a matrix $\mathbf{H} \in \mathbb{R}^{\bar{p} \times p}$ with rank p, such that the following equation is verified $\bar{\mathbf{y}}_t = \mathbf{H}\mathbf{y}_t \ \forall \mathbf{y_t} \in \mathbb{R}^p$. We set $\bar{\mathbf{C}} = \mathbf{HC}$. From now on, we assume that the system partitioning has been carried out in such a way that the following assumption holds.

Assumption 1: The pairs $(A^{[i]}, C^{[i]})$ are observable, for all $i = 1, ..., M$.

Notice that, neither Assumption 1 implies that the pair $(\mathbf{A}, \mathbf{C})$ is observable, nor observability of (1)-(2) implies Assumption 1.

We define n_i^o as the observability index of the pair $(A^{[i]}, C^{[i]})$[1] and $\bar{n}^o = \max_{i=1,\ldots,M}(n_i^o)$. Denote $\mathbf{A}^* = \text{diag}(A^{[1]}, \ldots, A^{[M]})$, $\tilde{\mathbf{A}} = \mathbf{A} - \mathbf{A}^*$, where $\tilde{\mathbf{A}}$ has the structure $\tilde{\mathbf{A}} = \left[(\tilde{A}^{[1]})^T \quad \ldots \quad (\tilde{A}^{[M]})^T\right]^T$, and $\tilde{A}^{[i]} \in \mathbb{R}^{n_i \times n}$. Furthermore $\mathbf{C}^* = \text{diag}(C^{[1]}, \ldots, C^{[M]})$, $\tilde{\mathbf{C}} = \bar{\mathbf{C}} - \mathbf{C}^*$, where $\tilde{\mathbf{C}}$ has the structure $\tilde{\mathbf{C}} = \left[(\tilde{C}^{[1]})^T \quad \ldots \quad (\tilde{C}^{[M]})^T\right]^T$, and $\tilde{C}^{[i]} \in \mathbb{R}^{p_i \times n}$.

The inputs $u_t^{[i],x}$ and $u_t^{[i],y}$ in (3) and (4) are computed according to the algebraic equations $u_t^{[i],x} = \tilde{A}^{[i]} \mathbf{x}_t$ and $u_t^{[i],y} = \tilde{C}^{[i]} \mathbf{x}_t$. We say that a system partition is *trivial* if $M = 1$.

[1] In view of Assumption 1, n_i^o is defined as the minimum value of N such that the matrix $\left[(C^{[i]})^T \quad \ldots \quad (C^{[i]}(A^{[i]})^{N-1})^T\right]^T$ has full column rank n_i

III. Three moving horizon partition-based algorithms

Our aim is to design, for each of the subsystems, an algorithm for computing a reliable estimate $\hat{x}_{t/t}^{[i]}$ of the subsystem's state $x_t^{[i]}$, based on the measurements $y_t^{[i]}$ and on the estimates of the crosstalk terms $u_t^{[i],x}$ and $u_t^{[i],y}$ provided by the estimators associated to the other subsystems. To this end, we propose three solutions, named PMHE1, PMHE2 and PMHE3.

A. Models for estimation and transmission of information

We denote with $\hat{x}_{t_1/t_2}^{[i]}$ the estimate of $x_{t_1}^{[i]}$ performed at time t_2 by subsystem i. Its error covariance matrix is denoted with $\Pi_{t_1/t_2}^{[i]}$ and we define

$$\hat{\mathbf{x}}_{t_1/t_2} = [(\hat{x}_{t_1/t_2}^{[1]})^T \ldots (\hat{x}_{t_1/t_2}^{[M]})^T]^T \tag{5a}$$

We approximate $\text{Var}(\mathbf{x}_{t_1} - \hat{\mathbf{x}}_{t_1/t_2})$ as

$$\mathbf{\Pi}_{t_1/t_2} = \text{diag}(\Pi_{t_1/t_2}^{[1]} \ldots \Pi_{t_1/t_2}^{[M]}) \tag{5b}$$

that corresponds to assume that the errors of different subsystems are uncorrelated. This approximation will allow decentralization of the centralized MHE problem. At time t the estimation model is, for $k = t-N, \ldots, t$

$$\hat{x}_{k+1}^{[i]} = A^{[i]} \hat{x}_k^{[i]} + \tilde{A}^{[i]} \tilde{\mathbf{x}}_{k/t-1} + \hat{w}_k^{[i]} \tag{6a}$$

$$y_k^{[i]} = C^{[i]} \hat{x}_k^{[i]} + \tilde{C}^{[i]} \tilde{\mathbf{x}}_{k/t-1} + \hat{v}_k^{[i]} \tag{6b}$$

and defines constraints of the PMHE estimation problem specified in the next section. In (6), $\tilde{\mathbf{x}}_{k/t-1} \in \mathbb{R}^n$ denotes estimates of the subsystem states available at time t, and can differ from $\hat{\mathbf{x}}_{k/t-1}$. Next we introduce two models for $\tilde{\mathbf{x}}_{k/t-1}$, $\hat{w}_k^{[i]}$ and $\hat{v}_k^{[i]}$, that are related to different communication protocols: the first one will be used in PMHE1, while the second one will be used in PMHE2 and PMHE3.

Model 1: the system partition induces an interconnected network of subsystems, which can be described by a directed graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$, where the nodes in $\mathscr{V}$ are the subsystems and the edge (j,i) in the set $\mathscr{E} \subseteq \mathscr{V} \times \mathscr{V}$ models that the j-th subsystem influences the dynamics of the i-th subsystem. Therefore, we assume that, at time t, if $(j,i) \in \mathscr{E}$, then $\hat{x}_{k/t-1}^{[j]}$ and $\Pi_{k/t-1}^{[j]}$ for $k = t-N, \ldots, t$ are transmitted to subsystem i. The noise terms $\hat{w}_k^{[i]}$ and $\hat{v}_k^{[i]}$ in (6) encompass both the noise appearing in the equations (3), (4) and the estimation error of the variables $u_t^{[i],x}$ and $u_t^{[i],y}$. Therefore their variance is given by

$$Var(\hat{w}_k^{[i]}) = \ Q^{[i]} + (\tilde{A}^{[i]}) \mathbf{\Pi}_{k/t-1} (\tilde{A}^{[i]})^T \tag{7a}$$

$$Var(\hat{v}_k^{[i]}) = \ R^{[i]} + (\tilde{C}^{[i]}) \mathbf{\Pi}_{k/t-1} (\tilde{C}^{[i]})^T \tag{7b}$$

Moreover, we set $\tilde{\mathbf{x}}_{k/t-1} = \hat{\mathbf{x}}_{k/t-1}$. Note that, in (6) and (7), the terms $\tilde{A}^{[i]} \hat{\mathbf{x}}_{k/t-1}$, $\tilde{C}^{[i]} \hat{\mathbf{x}}_{k/t-1}$, $Var(\hat{w}_k^{[i]})$ and $Var(\hat{v}_k^{[i]})$ depend only upon the quantities transmitted by the neighboring subsystems $j \in \mathscr{V}^{[i]} = \{j : (j,i) \in \mathscr{E}\}$.

Model 2: we assume an all-to-all communication, so that all the subsystems at time $t-1$ know the vector $\hat{\mathbf{x}}_{t-N/t-1}$ and, for PMHE2, the matrix $\mathbf{\Pi}_{t-N/t-1}$. Accordingly, at time t, the i-th subsystems estimation model, for $k=t-N,\cdots,t$, is (6), where $\tilde{\mathbf{x}}_{k/t-1}=\mathbf{A}^{k-(t-N)}\hat{\mathbf{x}}_{t-N/t-1}$, for $k=t-N,\dots,t$. The noise terms $\hat{w}_k^{[i]}$ and $\hat{v}_k^{[i]}$ encompass now also the uncertainty characterizing the terms $\tilde{A}^{[i]}\tilde{\mathbf{x}}_{k/t-1}$ and $\tilde{C}^{[i]}\tilde{\mathbf{x}}_{k/t-1}$, respectively, and hence their covariance is given by

$$Var(\hat{w}_k^{[i]}) = Q^{[i]} + \tilde{A}^{[i]}\mathbf{\Pi}^{ol}_{k/t-1}(\tilde{A}^{[i]})^T \quad (8a)$$

$$Var(\hat{v}_k^{[i]}) = R^{[i]} + (\tilde{C}^{[i]})\mathbf{\Pi}^{ol}_{k/t-1}(\tilde{C}^{[i]})^T \quad (8b)$$

where $\mathbf{\Pi}^{ol}_{t-N/t-1}=\mathbf{\Pi}_{t-N/t-1}$ and, for $k\geq t-N+1$

$$\mathbf{\Pi}^{ol}_{k/t-1} = \mathbf{A}^{k-(t-N)}\mathbf{\Pi}_{t-N/t-1}(\mathbf{A}^{k-(t-N)})^T + \sum_{i=t-N+1}^{k}\mathbf{A}^{i-(t-N+1)}\mathbf{Q}(\mathbf{A}^{i-(t-N+1)})^T \quad (8c)$$

B. The PMHE1 and PMHE2 estimation problems

For a given estimation horizon $N\geq 1$, each node $i\in\mathscr{V}$ at time t solves the constrained minimization problem MHE-i defined as

$$\min_{\hat{x}^{[i]}_{t-N},\{\hat{w}_k^{[i]}\}_{k=t-N}^{t-1}} J^{[i]}(t-N,t,\hat{x}^{[i]}_{t-N},\hat{w}^{[i]},\hat{v}^{[i]},\Gamma^{[i]}_{t-N}) \quad (9)$$

where, for brevity, $\hat{w}^{[i]}$ and $\hat{v}^{[i]}$ stand for $\{\hat{w}_k^{[i]}\}_{k=t-N}^{t-1}$ and $\{\hat{v}_k^{[i]}\}_{k=t-N+1}^{t}$, respectively, under the constraints

$$\begin{cases} \text{System (6) with transmission Model 1 for PMHE1} \\ \text{System (6) with transmission Model 2 for PMHE2} \end{cases} \quad (10a)$$

$$\hat{w}_k^{[i]}\in\mathbb{W}_i \text{ and } \hat{x}_k^{[i]}\in\mathbb{X}_i \quad (10b)$$

where $k=t-N,\dots,t$ and the local cost function $J^{[i]}$ is given by

$$J^{[i]}(t-N,t,\hat{x}^{[i]}_{t-N},\hat{w}^{[i]},\hat{v}^{[i]},\Gamma^{[i]}_{t-N}) = \tfrac{1}{2}(\textstyle\sum_{k=t-N}^{t}\|\hat{v}_k^{[i]}\|^2_{(R^{*[i]}_{k/t-1})^{-1}} + \textstyle\sum_{k=t-N}^{t-1}\|\hat{w}_k^{[i]}\|^2_{(Q^{[i]*}_{k/t-1})^{-1}}) + \Gamma^{[i]}_{t-N}(\hat{x}^{[i]}_{t-N};\hat{x}^{[i]}_{t-N/t-1}) \quad (11)$$

Let $\hat{x}^{[i]}_{t-N/t}$ and $\left\{\hat{w}^{[i]}_{k/t}\right\}_{k=t-N}^{t-1}$ be the optimizers to (9) and $\hat{x}^{[i]}_{k/t}$, $k=t-N,...,t$ the local state sequence stemming from $\hat{x}^{[i]}_{t-N/t}$ and $\left\{\hat{w}^{[i]}_{k/t}\right\}_{k=t-N}^{t-1}$. In (11), the function $\Gamma^{[i]}_{t-N}(\hat{x}^{[i]}_{t-N};\hat{x}^{[i]}_{t-N/t-1})$ is the so called *initial penalty*, defined as follows

$$\Gamma^{[i]}_{t-N}(\hat{x}^{[i]}_{t-N};\hat{x}^{[i]}_{t-N/t-1}) = \frac{1}{2}\|\hat{x}^{[i]}_{t-N}-\hat{x}^{[i]}_{t-N/t-1}\|^2_{(\Pi^{[i]}_{t-N/t-1})^{-1}}, \quad (12)$$

In (11) and hereafter, the notation $\|z\|^2_S$ stands for z^TSz, where S is a positive-semidefinite matrix. Notice that, in case of gaussian uncertainties, the problem (11) can be interpreted as a maximum *a posteriori* likelihood problem, see [8]. The positive definite symmetric matrix $\Pi^{[i]}_{t-N/t-1}$ appearing in (12) plays the role of a covariance matrix and is a design parameter whose choice is discussed in detail in the next section.

C. Computation of $\Pi^{[i]}_{t-N/t-1}$ for PMHE1 and PMHE2

We choose the matrix $\Pi^{[i]}_{t-N/t-1}$, $i\in\mathscr{V}$, as the result of one iteration of the difference Riccati equation associated to a Kalman filter for the system

$$\begin{cases} x^{[i]}_{t-N} &= A^{[i]}x^{[i]}_{t-N-1}+w^{[i]}_{t-N-1} \\ z^{[i]}_{t-N} &= \mathscr{O}^{[i]}_N x^{[i]}_{t-N}+\mathscr{C}^{[i]}_N X^{[t-N,t-1]/t-2}+V^{[i]}_{t-N} \end{cases}$$

where $X^{[t-N,t-1]/t-2}=[(\hat{\mathbf{x}}_{t-N/t-2})^T,\dots,(\hat{\mathbf{x}}_{t-1/t-2})^T]^T$ are inputs, $\mathscr{O}^{[i]}_{N+1}$ is the extended observability matrix of the pair $(A^{[i]},C^{[i]})$ defined as

$$\mathscr{O}^{[i]}_{N+1} = \begin{bmatrix}(C^{[i]})^T & \dots & (C^{[i]}(A^{[i]})^N)^T\end{bmatrix}^T \quad (13a)$$

and

$$\mathscr{C}^{[i]}_N = \begin{bmatrix} \tilde{C}^{[i]} & 0 & \dots & 0 \\ C^{[i]}\tilde{A}^{[i]} & \tilde{C}^{[i]} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ C^{[i]}(A^{[i]})^{N-2}\tilde{A}^{[i]} & C^{[i]}(A^{[i]})^{N-3}\tilde{A}^{[i]} & \dots & \tilde{C}^{[i]} \end{bmatrix} \quad (13b)$$

$$\mathscr{C}^{[i]}_{w,N} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ C^{[i]} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ C^{[i]}(A^{[i]})^{N-2} & C^{[i]}(A^{[i]})^{N-3} & \dots & C^{[i]} \end{bmatrix} \quad (13c)$$

$$\mathbf{R}^{*[i]}_{N/t-2} = \text{diag}\left(R^{*[i]}_{t-N/t-2},\dots,R^{*[i]}_{t-1/t-2}\right) \quad (13d)$$

$$\mathbf{Q}^{*[i]}_{N-1/t-2} = \text{diag}\left(Q^{*[i]}_{t-N/t-2},\dots,Q^{*[i]}_{t-2/t-2}\right) \quad (13e)$$

$$\text{Cov}[w^{[i]}_{t-N-1}] = Q^{*[i]}_{t-N-1/t-2} \quad (13f)$$

$$\text{Cov}[\bar{V}^i_t] = \tilde{\mathbf{R}}^{[i]}_{N/t-2} = \mathbf{R}^{*[i]}_{N/t-2} + \mathscr{C}^{[i]}_{w,N}\mathbf{Q}^{*[i]}_{N-1/t-2}(\mathscr{C}^{[i]}_{w,N})^T \quad (13g)$$

In the Kalman filter, the covariance of the estimate $\hat{x}^{[i]}_{t-N-1/t-2}$ is set as

$$\bar{\Pi}^{[i]}_{t-N-1/t-2} = \left((\Pi^{[i]}_{t-N-1/t-2})^{-1} + (C^{[i]})^T(R^{*[i]}_{t-N-1/t-2})^{-1}C^{[i]}\right)^{-1} \quad (14a)$$

As a result, we obtain the Riccati equation

$$\begin{aligned}\Pi^{[i]}_{t-N/t-1} =& A^{[i]}\bar{\Pi}^{[i]}_{t-N-1/t-2}(A^{[i]})^T + Q^{*[i]}_{t-N-1/t-2} + \\ & - A^{[i]}\bar{\Pi}^{[i]}_{t-N-1/t-2}(\mathscr{O}^{*[i]}_N)^T \times \\ & \times \left(\mathscr{O}^{[i]}_N\bar{\Pi}^{[i]}_{t-N-1/t-2}(\mathscr{O}^{[i]}_N)^T + \tilde{\mathbf{R}}^{[i]}_{N/t-2}\right)^{-1} \times \\ & \times \mathscr{O}^{[i]}_N\bar{\Pi}^{[i]}_{t-N-1/t-2}(A^{[i]})^T\end{aligned} \quad (14b)$$

Note that, for the computation of $\Pi^{[i]}_{t-N/t-1}$, the matrix update (14a) and (14b) is applied, and that $R^{*[i]}_{k/t-2}$ and $Q^{*[i]}_{k/t-2}$ ($k=t-N,\dots,t-1$) are computed as in (7) for PMHE1 [resp. (8) for PMHE2]. Therefore, such recursive equations have, as input terms

- the error covariance matrices $\Pi^{[j]}_{k/t-2}$ ($k=t-N,\dots,t-1$) of the neighbors to subsystem i, *i.e.*, the subsystems

Algorithm	Type of transmission	Order of the optimization problem	Transmitted information
Centralized MHE	NO transmission	$n\times(\bar{n}^o+1)$	
PMHE1	Neighbor-to-neighbor	$n_i\times(\bar{n}^o+1)$	$\hat{x}^{[j]}_{k/t-1},\Pi^{[j]}_{k/t-1},k=t-N,\dots,t$
PMHE2	All-to-all	$n_i\times(\bar{n}^o+1)$	$\hat{x}^{[j]}_{t-N/t-1},\Pi^{[j]}_{t-N/t-1}$
PMHE3	All to all	n_i	$\hat{x}^{[j]}_{t-N/t-1}$

TABLE I

Comparison of PMHE1, PMHE2, PMHE3 and centralized MHE in terms of transmission requirements and computational load.

indexed by the set $\mathscr{V}_i=\{j:(j,i)\in\mathscr{E}\}$ in case of PMHE1;

- $\mathbf{\Pi}_{t-N/t-1}$, used in (8), in case of PMHE2.

The positive feedback effect emerging from this interaction might cause unboundedness of the sequence $\Pi^{[i]}_{t-N/t-1}$, $\forall i\in\mathscr{V}$, which must be avoided, in order to guarantee boundedness of the weighting matrices $Q^{*[i]}_{k/t-1}$ and $R^{*[i]}_{k/t-1}$ and the applicability of the proposed PMHE1 and PMHE2 algorithms. Some solutions are proposed in [6] for guaranteeing boundedness of $\Pi^{[i]}_{k/t-1}$, at the price of suboptimality in noise filtering.

D. The PMHE3 estimation problem

For a given estimation horizon $N\geq 1$, each node $i\in\mathscr{V}$ at time t solves the optimization problem

$$\min_{\hat{x}^{[i]}_{t-N}} J^{[i]}_3(t-N,t,\hat{x}^{[i]}_{t-N},\hat{v}^{[i]},\Gamma^{[i]}_{3,t-N}) \quad (15)$$

under the constraints

$$\text{system (6) with transmission Model 2} \quad (16a)$$
$$w^{[i]}_k=0 \text{ for } k=t-N,\dots,t-1 \quad (16b)$$
$$\hat{x}^{[i]}_k\in\mathbb{X}_i \text{ for } k=t-N,\dots,t \quad (16c)$$

and the local cost function $J^{[i]}_3$ is given by

$$J^{[i]}_3(t-N,t,\hat{v}^{[i]},\hat{x}^{[i]}_{t-N})=\frac{1}{2}\sum_{k=t-N}^{t}\|\hat{v}^{[i]}_k\|^2+\Gamma^{[i]}_{3,t-N}(\hat{x}^{[i]}_{t-N};\hat{x}^{[i]}_{t-N/t-1}) \quad (17)$$

The term $\Gamma^{[i]}_{3,t-N}(\hat{x}^{[i]}_{t-N};\hat{x}^{[i]}_{t-N/t-1})$ is the initial penalty defined as

$$\Gamma^{[i]}_{3,t-N}(\hat{x}^{[i]}_{t-N};\hat{x}^{[i]}_{t-N/t-1})=\frac{\mu}{2}\|\hat{x}^{[i]}_{t-N}-\hat{x}^{[i]}_{t-N/t-1}\|^2$$

where $\mu\geq 0$. Moreover, notice that $\hat{x}^{[i]}_{t-N/t-1}$ is computed as

$$\hat{x}^{[i]}_{t-N/t-1}=A^{[i]}\hat{x}^{[i]}_{t-N-1/t-1}+\tilde{A}^{[i]}\hat{\mathbf{x}}_{t-N-1/t-1}$$

E. Communication requirements and computational load

The three solutions proposed in this section have different features in terms of communication requirements among subsystems, accuracy and computational complexity (see Table I). More specifically, PMHE1 relies on a partially connected communication graph in the sense that subsystems exploit a communication network where links are present only if subsystem dynamics are coupled. Algorithms PMHE2 and PMHE3 assume an all-to-all communication but a reduced amount of information is transmitted over each communication channel. The main difference between PMHE1 and PMHE2 consists in the type of communication required among subsystems, and on how the estimates of $u^{[i],x}_t$ and $u^{[i],y}_t$ are used. While in PMHE1 and PMHE2 the transmitted information amounts to state estimates and estimation error covariances, in PMHE3 no information on the noise variances is required and the weights on the different components of the cost functions are constant, allowing for a significant reduction in terms of transmission and computational load.

IV. Convergence properties of the proposed estimators

The main purpose of this section is to extend the convergence results of [13] and [1] to the proposed PMHE methods in presence of constraints. Similarly to [13], the analysis is conducted in a deterministic setting.

Definition 1: Let Σ be system (1) with $\mathbf{w}_t=0$ and denote by $\mathbf{x}_\Sigma(t,\mathbf{x}_0)$ the state reached by Σ at time t starting from initial condition $\mathbf{x}_0$. Assume that the trajectory $\mathbf{x}_\Sigma(t,\mathbf{x}_0)$ is feasible, *i.e.,* $\mathbf{x}_\Sigma(t,\mathbf{x}_0)\in\mathbb{X}$ for all t. PMHE is *convergent* if $\|\hat{\mathbf{x}}_{t/t}-\mathbf{x}_\Sigma(t,\mathbf{x}_0)\|\overset{t\to\infty}{\longrightarrow}0$. □

Note that, as in [13], convergence is defined assuming that the model generating the data is noiseless, but the possible presence of noise is taken into account in the state estimation algorithm.

The estimation error is defined as $\varepsilon_{k_1/k_2}=\mathbf{x}_\Sigma(k_1,\mathbf{x}_0)-\hat{\mathbf{x}}_{k_1/k_2}$. Let $\mathscr{O}^*_{N+1}$ and $\mathscr{O}_{N+1}$ be the extended observability matrices, defined as in (13a), of the pairs $(\mathbf{A}^*,\mathbf{C}^*)$ and $(\mathbf{A},\bar{\mathbf{C}})$, respectively. Note that, by construction, $\mathscr{O}^*_{N+1}=\text{diag}\left(\mathscr{O}^{[1]}_{N+1},\dots,\mathscr{O}^{[M]}_{N+1}\right)\in\mathbb{R}^{(N+1)\bar{p}\times n}$. We denote by $f_{min}=\sigma_{min}(\mathscr{O}^*_{N+1})$ and $f_{max}=\sigma_{max}(\mathscr{O}^*_{N+1})$, the minimum and the maximum singular value of $\mathscr{O}^*_{N+1}$, respectively. By Assumption 1, if $N\geq\bar{n}^o-1$, then $\text{rank}(\mathscr{O}^{[i]}_{N+1})=n_i$ for all $i\in\mathscr{V}$. From this it follows that $\text{rank}(\mathscr{O}^*_{N+1})=n$, and therefore $f_{min}>0$. Furthermore, define $\Delta_f=\|\mathscr{O}^*_{N+1}-\mathscr{O}_{N+1}\|_2$, $\kappa=\|\mathbf{A}\|_2$, and $\kappa^*=\|\mathbf{A}^*\|_2$.

Then, the following results can be established. Their proof is provided in [6].

Lemma 1: If matrices $\Pi^{[i]}_{t-N/t-1}$ are computed as in Section III-C and $N\geq\max\{\bar{n}^o-1,1\}$, then the dynamics of the state estimation error generated by PMHE1 is given by

$$\mathscr{O}^*_{N+1}\varepsilon_{t-N/t}=-\mathscr{C}_{N+1}E_{t/t-1}+\alpha^1_t \quad (18a)$$
$$E_{t+1/t}-M_1\varepsilon_{t-N/t}=M_2E_{t/t-1}+\alpha^2_t \quad (18b)$$

where $E_{k_1/k_2} = \begin{bmatrix} \varepsilon^T_{k_1-N/k_2} & \dots & \varepsilon^T_{k_1/k_2} \end{bmatrix}^T$ and

$$\mathscr{C}_{N+1} = \begin{bmatrix} \tilde{\mathbf{C}} & 0 & \dots & 0 \\ \mathbf{C}^*\tilde{\mathbf{A}} & \tilde{\mathbf{C}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}^*(\mathbf{A}^*)^{N-1}\tilde{\mathbf{A}} & \mathbf{C}^*(\mathbf{A}^*)^{N-2}\tilde{\mathbf{A}} & \dots & \tilde{\mathbf{C}} \end{bmatrix} \quad (19a)$$

$$M_1 = \begin{bmatrix} \mathbf{A}^* \\ \vdots \\ (\mathbf{A}^*)^N \\ (\mathbf{A}^*)^{N+1} \end{bmatrix}, \; M_2 = \begin{bmatrix} \tilde{\mathbf{A}} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbf{A}^*)^{N-1}\tilde{\mathbf{A}} & (\mathbf{A}^*)^{N-2}\tilde{\mathbf{A}} & \dots & 0 \\ (\mathbf{A}^*)^{N}\tilde{\mathbf{A}} & (\mathbf{A}^*)^{N-1}\tilde{\mathbf{A}} & \dots & \tilde{\mathbf{A}} \end{bmatrix} \quad (19b)$$

and α^j_t are asymptotically vanishing terms, i.e. $\|\alpha^j_t\| \xrightarrow{t\to\infty} 0$, $j = 1, 2$. □

Lemma 2: If matrices $\Pi^{[i]}_{t-N/t-1}$ are computed as in Section III-C and $N \geq \max\{\bar{n}^o - 1, 1\}$, then the dynamics of the state estimation error generated by PMHE2 is given by

$$\mathscr{O}^*_{N+1}\varepsilon_{t-N/t} = \left(\mathscr{O}^*_{N+1} - \mathscr{O}_{N+1}\right)\mathbf{A}\varepsilon_{t-N-1/t-1} + \alpha_t \quad (20a)$$

where α_t is an asymptotically vanishing term, i.e. $\|\alpha_t\| \xrightarrow{t\to\infty} 0$. □

Lemma 3: Assume that $\mathbf{A}^*$ is non singular, and that one of the following conditions holds: (a) $\kappa^* \leq 1$, (b) $\kappa^* > 1$ and $\mu < \mu_{max}$, where $\mu_{max} = \frac{f^2_{min}}{(\kappa^*)^2 - 1}$. If $N \geq \max\{\bar{n}^o - 1, 1\}$, then the dynamics of the state estimation error generated by PMHE3 obeys to (20). □

From these lemmas, sufficient conditions for the convergence of the PMHE algorithms are given in the next theorem.

Theorem 1:

I) Under the assumptions of Lemma 1, PMHE1 is convergent if the matrix Φ_1 is Schur, where

$$\Phi_1 = M_2 - M_1\left((\mathscr{O}^*_{N+1})^T\mathscr{O}^*_{N+1}\right)^{-1}(\mathscr{O}^*_{N+1})^T\mathscr{C}_{N+1}$$

II) Under the assumptions of Lemma 2, PMHE2 is convergent if the matrix Φ_2 is Schur, where

$$\Phi_2 = \left((\mathscr{O}^*_{N+1})^T\mathscr{O}^*_{N+1}\right)^{-1}(\mathscr{O}^*_{N+1})^T\left(\mathscr{O}^*_{N+1} - \mathscr{O}_{N+1}\right)\mathbf{A} \quad (21)$$

III) Under the assumptions of Lemma 3, PMHE3 is convergent if the matrix Φ_2 is Schur.

V. Example: a compartmental system

Consider the interconnected system reported in Figure 1-*A*. The subsystems 1, …, 4 are third order compartmental subsystems, whose structure is depicted in Figure 1-*B*. If subsystem i has m inputs and p outputs, its discrete-time dynamical model is defined by

$$A^{[i]} = \begin{bmatrix} 1-k_{12} & k_{21} & k_{31} \\ k_{12} & 1-(k_{21}+k_{23}+pk_{2y}) & 0 \\ 0 & k_{23} & 1-k_{13} \end{bmatrix}$$

$$B^{[i]} = \underbrace{\begin{bmatrix} 1 & \dots & 1 \\ 0 & \dots & 0 \\ 0 & \dots & 0 \end{bmatrix}}_{m \text{ columns}}, \; C^{[i]*} = \left.\begin{bmatrix} 0 & k_{2y} & 0 \\ \vdots & \vdots & \vdots \\ 0 & k_{2y} & 0 \end{bmatrix}\right\} p \text{ rows}$$

We chose $k_{ij} = 0.1$ for all i and j, $k_{2y} = 0.1$, and we introduce the vectors $b = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$ and $c = \begin{bmatrix} 0 & k_{2y} & 0 \end{bmatrix}$. If we connect the 4 subsystems according to the scheme in Figure 1-*A* we obtain a 12-states system with the structure (1) and (2), given by

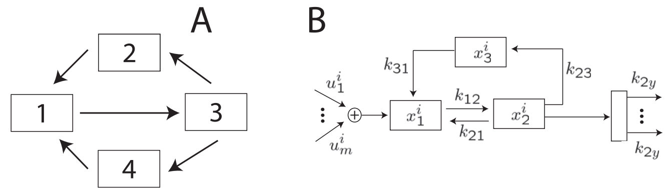

Fig. 1. Scheme of the compartmental system in the example. A: connections between subsystems 1, …, 4. B: general structure of the subsystems 1, …, 4.

$$\mathbf{A} = \begin{bmatrix} A^{[1]} & bc & \mathbf{O}_{3\times3} & bc \\ \mathbf{O}_{3\times3} & A^{[2]} & bc & \mathbf{O}_{3\times3} \\ bc & \mathbf{O}_{3\times3} & A^{[3]} & \mathbf{O}_{3\times3} \\ \mathbf{O}_{3\times3} & \mathbf{O}_{3\times3} & bc & A^{[4]} \end{bmatrix}$$

and $\mathbf{C} = \mathrm{diag}(c, c, c, c)$. Note that the spectral radius of $\mathbf{A}$ is 1. We assume that states of each subsystem are affected by leakages represented by additive negative noise terms $\mathbf{w}_k$. We take $\mathbf{w}_k = \max(-1, -|\mathbf{e}_k|)$, for all k, where $\mathbf{e}_k$ is a white noise signal with zero mean and $\mathbf{Q} = var(\mathbf{e}_k) = \mathrm{diag}(1, \varepsilon, \varepsilon, 1, \varepsilon, \varepsilon, 1, \varepsilon, \varepsilon, 1, \varepsilon, \varepsilon)$, where $\varepsilon = 10^{-8}$. Therefore, the first state of each subsystem is affected by leakage more severely than the other states. We assume white measurement noise with covariance $\mathbf{R} = 0.01I_{12}$. Since the states represent masses in the compartments, they are constrained to be non negative. Furthermore, we take $\mathbf{\Pi}_0 = 340I_{12}$.

Next we compare the PMHE1, PMHE2, PMHE3 strategies with a centralized MHE estimator. For the design of PMHE3, we compute $\kappa^* = 0.9913 < 1$ and, by Lemma 3, all $\mu > 0$ guarantee convergence of the estimates. We choose $\mu = 0.001$. The convergence properties of PMHE estimators can be proved using Theorem 1. In order to guarantee the applicability of the four estimators, the estimation horizon is set as $N = 3$ in all the PMHE schemes (to satisfy the assumptions of Lemmas 1, 2 and 3) as well as in the centralized MHE. In Fig. 2 we compare the estimated and real state trajectories. We have also explored the effect of the variation of the estimation horizon N on the estimation performances and on the computational burden through simulations. The root mean square error for $t \in [15, 45]$ (*i.e.*, neglecting the initial transients) and the time required to run the estimation algorithms, for $N = 3, 7, 10$, are reported in Table II.

Interestingly, the time required for each node to perform PMHE1, PMHE2 and PMHE3, is reduced with respect to the time required to perform centralized MHE, at the price of obtaining suboptimal estimations in terms of noise rejection. Although as N increases a larger set of data is used in the optimization problem, this does not lead a significant improvement of the accuracy of the results. On the other hand, an increase in N leads to a significant grow in computational (and transmission) burden (i.e. $T_c = 3.8$s if $N = 3$, $T_c = 11$s

	PMHE1			PMHE2			PMHE3		
N	3	7	10	3	7	10	3	7	10
RMSE (rel)	$1.2J_c$	$1.3J_c$	$1.2J_c$	$1.2J_c$	$1.3J_c$	$1.3J_c$	$1.7J_c$	$1.7J_c$	$1.6J_c$
RMSE (abs)	4.43	4.36	4.38	4.42	4.47	4.42	6.1	5.84	5.76
Req. time (rel)	$0.79T_c$	$0.34T_c$	$0.2T_c$	$0.8T_c$	$0.34T_c$	$0.2T_c$	$0.4T_c$	$0.13T_c$	$0.07T_c$

TABLE II

Comparison of the root mean square error (RMSE) $\frac{1}{31}\sum_{k=15}^{45}\|\hat{x}_{k/k}-x_k\|_2^2$ and of the time required (for each subsystem) to perform the proposed estimates. J_c and T_c denote the RMSE of the centralized MHE estimation error and the time required to perform the centralized MHE, respectively.

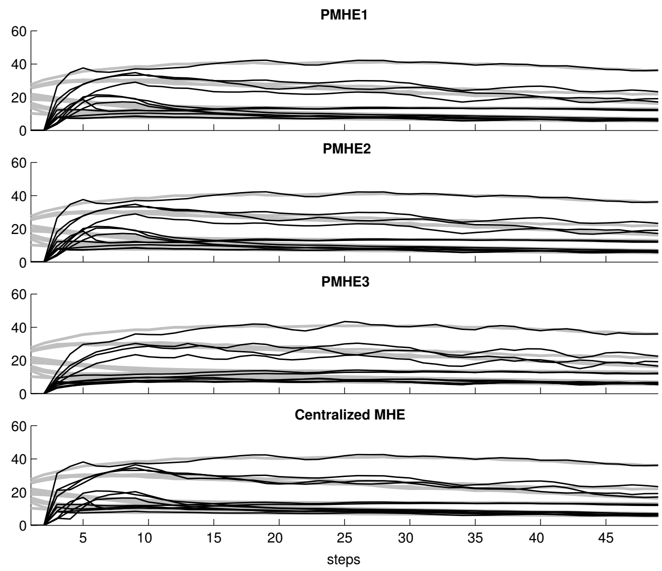

Fig. 2. Real trajectories $\mathbf{x}_t$ (grey lines) and estimated trajectories $\hat{\mathbf{x}}_{t/t}$ (black lines) for the different estimation schemes.

if $N = 7$ and $T_c = 21.9$s if $N = 10$).

VI. Conclusions

In this paper we have proposed three decentralized MHE schemes for partitioned large-scale systems. Sufficient conditions for convergence of the state-estimate to the true state have been derived. Algorithms PMHE2 and PMHE3 require an all-to-all communication network and future research will focus on methods for weakening this assumption. A promising research direction is to merge PMHE with ideas of distributed MHE [4], [5] in order to generalize the proposed schemes to the case of subsystems with overlapping states and guarantee that estimates of states that are shared by some subsystems converge to a common value. Further studies are also needed for designing PMHE schemes capable to cope with non-idealties in the communication network such as quantization and transmission delays.

References

[1] A. Alessandri, M. Baglietto, and G. Battistelli. Receding-horizon estimation for discrete-time linear systems. *IEEE Trans. on Automatic Control*, 48(3):473 – 478, 2003.

[2] A. Alessandri, M. Baglietto, and G. Battistelli. Moving-horizon state estimation for nonlinear discrete-time systems: new stability results and approximation schemes. *Automatica*, 44:1753 – 1765, 2008.

[3] A. Alessandri, M. Baglietto, T. Parisini, and R. Zoppoli. A neural state estimator with bounded errors for nonlinear systems. *IEEE Transactions on Automatic Control*, 44(11):2028 – 2042, 1999.

[4] M. Farina, G. Ferrari-Trecate, and R. Scattolini. Distributed moving horizon estimation for sensor networks. In *Proceedings of the 1st IFAC Workshop on Estimation and Control of Networked Systems*, pages 126–131, 2009.

[5] M. Farina, G. Ferrari-Trecate, and R. Scattolini. A moving horizon scheme for distributed state estimation. In *Proceedings of the 48th Conference on Decision and Control, Shangai, China*, pages 1818–1823, 2009.

[6] M. Farina, G. Ferrari-Trecate, and R. Scattolini. Moving horizon state estimation of large-scale constrained partitioned systems. *Automatica. In press. DOI: 10.1016/j.automatica.2010.02.010.*, 2010.

[7] G. Ferrari-Trecate, D. Mignone, and M. Morari. Moving horizon estimation for hybrid systems. *Automatic Control, IEEE Trans. on*, 47(10):1663 – 1676, 2002.

[8] G.C. Goodwin, M.M. Seron, and J.A. De Doná. *Constrained Control and Estimation*. Springer, New Jersey, 2005.

[9] H.R. Hashemipour, S. Roy, and A.J. Laub. Decentralized structures for parallel Kalman filtering. *Automatic Control, IEEE Trans. on*, 33(1):88 – 94, Jan. 1988.

[10] U. A. Khan and J. M. F. Moura. Distributing the Kalman filter for large-scale systems. *Signal Processing, IEEE Trans. on*, 56(10):4919 – 4935, Oct. 2008.

[11] A. G. O. Mutambara. *Decentralized estimation and control for multisensor systems*. CRC Press, 1998.

[12] B.S. Rao and H.F. Durrant-Whyte. Fully decentralised algorithm for multisensor Kalman filtering. In *Control Theory and Applications, IEE Proceedings D*, volume 138, pages 413 – 420, Sept. 1991.

[13] C. V. Rao, J.B. Rawlings, and J.H. Lee. Constrained linear state estimation - a moving horizon approach. *Automatica*, 37:1619–1628, 2001.

[14] C. V. Rao, J.B. Rawlings, and D.Q Mayne. Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE Trans. on Automatic Control*, 48(2):246–258, 2003.

[15] C.V. Rao. *Moving Horizon Strategies for the Constrained Monitoring and Control of Nonlinear Discrete-Time Systems*. PhD thesis, University of Winsconsin-Madison, 2000.

[16] Jr. Sandell, N., P. Varaiya, M. Athans, and M. Safonov. Survey of decentralized control methods for large scale systems. *Automatic Control, IEEE Trans. on*, 23(2):108 – 128, Apr 1978.

[17] S.S. Stanković, M.S. Stanković, and D.M. Stipanović. Consensus based overlapping decentralized estimation with missing observations and communication faults. *Automatica*, 45:1397 – 1406, 2009.

[18] S.S. Stanković, M.S. Stanković, and D.M. Stipanović. Consensus based overlapping decentralized estimator. *Automatic Control, IEEE Trans. on*, 54(2):410 – 415, Feb. 2009.

[19] R. Vadigepalli and III Doyle, F.J. A distributed state estimation and control algorithm for plantwide processes. *Control Systems Technology, IEEE Trans. on*, 11(1):119 – 127, Jan. 2003.

[20] D. D. Šiljac. *Large-scale dynamic systems. Stability and structure.* North Holland, 1978.