A Decentralized Optimal LQ State Observer Based On An Augmented Lagrangian Approach

Didier Georges, Gildas Besancon, Jean-François Dulhoste

A Decentralized Optimal LQ State Observer Based On An Augmented Lagrangian Approach ⋆

Didier Georges [a], Gildas Besançon [a,b], Jean-François Dulhoste [c]

[a] *Grenoble INP / Univ. Grenoble 1 / Univ. Grenoble 3 / CNRS, GIPSA-lab, 38402, Saint Martin d'Hères, France,* {*didier.georges, gildas.besancon*}*@grenoble-inp.fr*

[b] *Institut Universitaire de France*

[c] *Universidad de los Andes, Mérida, Venezuela,* *djean@ula.ve*

Abstract

This paper is devoted to the design of a decentralized optimal batch LQ state observer for state estimation of large-scale interconnected systems, well suited for implementation on a sensor network. The here-proposed approach relies on both the use of an augmented Lagrangian formulation and a price-decomposition-coordination algorithm. The state estimation of an open-channel hydraulic system illustrates the effectiveness of this approach and is used to provide a comparison with alternative methods.

Key words: Decentralized optimal LQ state observer, price decomposition-coordination, augmented Lagrangian.

1 INTRODUCTION

The design of decentralized or distributed state estimation or control algorithms is an important research topic which has attracted a great interest for more than 30 years [1–11]. On the one hand, estimation of large-scale complex systems is of the greatest importance for monitoring applications in various engineering fields, such as power grids, road traffic networks, environmental systems (open-channel hydraulic systems, water supply networks, weather, ...). On the other hand, sensor networks [12] are now recognized for being well suited for measurement, monitoring, tracking of distributed physical phenomena, such as environmental phenomena (weather, seismic events, wildfires, air / soil / river pollution [13], sound or vibration monitoring ...) or complex industrial systems. Sensor networks are defined as a collection of embedded sensors with communication capabilities. Large scale deployment of such ad hoc networks mainly relies on the availability of cheap embedded sensors. For that reason, it is necessary to consider algorithmic architectures capable of performing low complexity computations distributed on each sensor node. Here the attention is paid to sensor network-based monitoring applications for large-scale systems, in which local measurements are made and a local state estimation with low computational complexity is performed by each sensor agent while some computation results are provided to the other agents connected to it. Furthermore a systematic way to manage decentralized state estimation through the application of dual optimization theory and the deterministic interpretation of the continuous-time Kalman filter as described in [18] is provided. Unlike the approaches in [1,9], the proposed approach provides an exact decentralized solution to the centralized Kalman filter and does not require strong structural properties of the system. The here-proposed approach is related to some previous works in [8], where dual decomposition is used. However the authors in [8] only consider the case of coupling of the subsystems via the measurement outputs. The here-proposed approach also differs from approaches based on consensus (see [5,4]), since the local systems are never some overlapping subsystems. It can be compared to the approach proposed in [6] since decentralization leads to the solution of distributed small-size moving horizon estimators without overlapping. However to the best of our knowledge, such an augmented

⋆ This work was partly supported by France-Venezuela ECOS program V11M02 / PI-2010000303. This paper was not presented at any IFAC meeting. Corresponding author D. Georges.

Email address: didier.georges@grenoble-inp.fr (Didier Georges).

Lagrangian formulation is used for the first time. It will be also shown that the proposed observer algorithm can be used to derive a moving horizon state observer. The paper is now organized as follows: In section II, the state estimation problem is stated. The decentralized design, based on both the use of an augmented Lagrangian formulation and a relaxation technique, together with some backgrounds on dual optimization, is then presented in section III. Section IV will be devoted to the derivation of a moving horizon observer based on the decentralized algorithm. Section V proposes an environmental state estimation application that illustrates the effectiveness of the approach. In this section, the proposed approach will be also compared to alternative approaches. Finally, the paper ends with some conclusions.

2 PROBLEM STATEMENT

We consider the problem of designing a decentralized state observer for large-scale interconnected linear systems defined by:

$$\dot{x}_i(t) = A_i x_i(t) + B_i u_i(t) + D_i \sum_{j=1}^{N} F_{ij} x_j(t), \quad (1)$$

$$y_i(t) = C_i x_i(t),\ i = 1, ..., N, \quad (2)$$

where the full state vector $x \in R^n$ and the full input vector $u \in R^m$ are partitioned as $x = (x_1^T, x_2^T, ..., x_N^T)^T$ and $u = (u_1^T, u_2^T, ..., u_N^T)^T$ respectively, where $x_i \in R^{n_i}$ and $u_i \in R^{m_i}$, and each subsystem admits a measurement output $y_i \in R^{p_i}$. The overall system is assumed to be observable.

System (1)-(2) may be equivalently defined as the following algebraic-differential system:

$$\dot{x}_i(t) = A_i x_i(t) + B_i u_i(t) + D_i s_i(t), \quad (3)$$

$$s_i(t) - \sum_{j=1}^{N} F_{ij} x_j(t) = 0, \quad (4)$$

$$y_i(t) = C_i x_i(t),\ i = 1, ..., N, \quad (5)$$

where the s_i's represent some slack variables. We seek to derive an optimal Linear Quadratic (LQ) state observer with a finite horizon by solving the following optimal output tracking problem, based on the knowledge of $y_i(t), u_i(t), \forall t \in I = [0, T]$:

$$\min_{v_i(t), \hat{s}_i(t), t \in I} \frac{1}{2} \sum_{i=1}^{N} \int_0^T \{\|\hat{y}_i(t) - y_i(t)\|^2_{R_i^{-1}} + \|v_i(t)\|^2_{Q_i^{-1}}\} dt + \frac{1}{2} \sum_{i=1}^{N} \|\hat{x}_i(0) - \underline{x}_i\|^2_{M_i^{-1}} \quad (6)$$

$$s.t.\ \dot{\hat{x}}_i(t) = A_i \hat{x}_i(t) + B_i u_i(t) i + E_i v_i(t) + D_i \hat{s}_i(t),$$

$$\hat{y}_i(t) = C_i \hat{x}_i(t),$$

$$\hat{s}_i(t) = \sum_{j=1}^{N} F_{ij} \hat{x}_j(t),$$

where Q_i, R_i and M_i, $i = 1, ..., N$ are some symmetric positive-definite matrices of adequate dimensions, and E_i's are also matrices of appropriate dimensions. $\underline{x}_i$ is the a priori most likely value of $x_i(0)$. This problem is an equivalent formulation to the continuous Kalman filter design problem which seeks the minimum variance state estimate of the overall state $\hat{x}$ of the system (see [18,20]), where Q_i is interpreted as the covariance matrix of a zero mean gaussian noise affecting the state equation of x_i through matrix E_i, and R_i is the covariance of a zero mean gaussian noise affecting the measurement output y_i. M_i is the covariance matrix of random initial state $x_i(0)$.

3 STATE OBSERVER DESIGN

The optimal output tracking problem (6) is closely related to the classical optimal trajectory tracking problem, except that the cost penalizes the initial states rather than the terminal states. For that reason, it appears to be very convenient to reverse time by considering the change $(t \to T - t)$ and interval $I' = [T, 0]$, and thus turning the optimal output tracking problem into:

$$\min_{\tilde{v}_i(t), \tilde{s}_i(t), i \in I'} \frac{1}{2} \sum_{i=1}^{N} \int_0^T \{\|C_i \tilde{x}_i(t) - \tilde{y}_i(t)\|^2_{R_i^{-1}} + \|\tilde{v}_i(t)\|^2_{Q_i^{-1}}\} dt + \frac{1}{2} \sum_{i=1}^{N} \|\tilde{x}_i(T) - \underline{x}_i\|^2_{M_i^{-1}}$$

$$s.t.$$

$$\dot{\tilde{x}}_i(t) = -A_i \tilde{x}_i(t) - B_i \tilde{u}_i(t) - E_i \tilde{v}_i(t) - D_i \tilde{s}_i(t),$$

$$\tilde{s}_i(t) = \sum_{j=1}^{N} F_{ij} \tilde{x}_j(t), \quad (7)$$

where $\tilde{x}_i(t) = \hat{x}_i(T - t)$,$\tilde{u}_i(t) = u_i(T - t)$, $\tilde{v}_i(t) = v_i(T - t)$ and $\tilde{s}_i(t) = \hat{s}_i(T - t)$. Let us now consider an *augmented Lagrangian* formulation of the problem as follows (in the spirit of [14,15] for instance):

$$L_c(\tilde{v}, \tilde{s}, \tilde{\mu}) = \frac{1}{2} \sum_{i=1}^{N} \int_0^T \{\|C_i \tilde{x}_i(t) - \tilde{y}_i(t)\|^2_{R_i^{-1}} + \|\tilde{v}_i(t)\|^2_{Q_i^{-1}}\} dt + \frac{1}{2} \sum_{i=1}^{N} \|\tilde{x}_i(T) - \underline{x}_i\|^2_{M_i^{-1}}$$

$$+ \sum_{i=1}^{N} \int_0^T < \tilde{\mu}_i(t) + \frac{c}{2} [\tilde{s}_i(t) - \sum_{j=1}^{N} F_{ij} \tilde{x}_j(t)],$$

$$\tilde{s}_i(t) - \sum_{j=1}^{N} F_{ij}\tilde{x}_j(t) > dt \qquad (8)$$

s.t.

$\dot{\tilde{x}}_i(t\,) = -A_i\tilde{x}_i(t) - B_i\tilde{u}_i(t) - E_i\tilde{v}_i(t) - D_i\tilde{s}_i(t),$

where $c > 0$ is the coefficient of the augmented Lagrangian, which should be chosen large enough to ensure the existence of a saddle-point. $\tilde{v}$, $\tilde{s}$, $\tilde{\mu}$ denote the vectors of all the $\tilde{v}_i$'s, $\tilde{s}_i$'s and $\tilde{\mu}_i$'s, respectively. $< .,. >$ represents the usual scalar product. μ is the vector of Lagrange multipliers associated to the interconnection constraints. *The interest of considering an augmented Lagrangian rather than the ordinary one may be found in the fact the problem is not strongly convex with respect to the slack variables* s_i. The augmented Lagrangian formulation may be viewed as a regularization technique which strongly convexifies the problem and therefore ensures the existence of a saddle-point of the constrained problem for non-strongly convex problems. *As a consequence, the relaxation algorithms, such as Uzawa algorithm [21], will converge by using a constant step gradient method, rather than a gradient method with "small steps" of type* σ [1] *, which tend to zero as the number of dual iterations tends to infinity (see [14])*, and therefore, cannot be easily handled in practice. On the basis of the augmented Lagrangian defined by (9), the problem consists in finding a saddle-point ($\tilde{v}$, $\tilde{s}$, $\tilde{\mu}$) of L_c, defined by

$$\min_{\tilde{v},\tilde{s}} \max_{\tilde{\mu}} L_c(\tilde{v},\tilde{s},\tilde{\mu}) = \max_{\tilde{\mu}} \min_{\tilde{v},\tilde{s}} L_c(\tilde{v},\tilde{s},\tilde{\mu}) \qquad (9)$$

where $w(\tilde{\mu}) = min_{\tilde{v},\tilde{s}} L_c(\tilde{v},\tilde{s},\tilde{\mu})$ is the so-called dual function, which has to be maximized.

3.1 Some background on dual optimization and augmented Lagrangians

In this paper, we consider iterative computation of a saddle-point solution of (9) by using a relaxation algorithm of the family of "Uzawa algorithms". In order to illustrate such an algorithm, let us consider the following constrained "separable" optimization problem:

$$\min_{u} J(u) = \min_{u_i} \sum_{i=1}^{N} J_i(u_i) \qquad (10)$$

$$s.t.\ \theta(u) = \sum_{i=1}^{N} \theta_i(u_i) = 0. \qquad (11)$$

and *the ordinary Lagrangian*

$$L(u,\mu) = J(u) + < \mu, \theta(u) > = \sum_{i=1}^{N} \{J_i(u_i) + \mu\theta_i(u_i)\}$$

where J is a convex functional defined on a Hilbert space $\mathcal{U} = \mathcal{U}_1 \times ... \times \mathcal{U}_N$ and the θ_i's are some vectorial mapping from $\mathcal{U}_i$ to $\mathcal{C}$. $\theta(u) = 0$ means that $\theta(u) \in -C = -\{0\}$, where C is the closed convex cone in $\mathcal{C}$ reduced to zero. In this case, the Uzawa algorithm becomes the so-called "price-decomposition-coordination algorithm" (PDCA) defined by:

(1) Initialization ($k = 0$): Choose u_i^0 and μ_i^0.
(2) At iteration k: Compute N independent subproblems: $u_i^{k+1} = arg\min_{u_i} J_i(u_i) + < \mu^k, \theta_i(u_i) >$, $i = 1, ..., N$.
(3) Update μ: $\mu^{k+1} = \mu^k + \rho^k \sum_{i=1}^{N} \theta_i(u_i^{k+1})$.
(4) if $\|\mu^{k+1} - \mu^k\| < \epsilon_c$, then stop, else $k+1 \to k$ and go to step (2).

At step (2), the problem splits into N subproblems (decomposition procedure), which may be solved in parallel with less complexity. This interesting feature was well exploited for computation of large-scale optimization problems, for which existence of a saddle-point is ensured (otherwise convergence is not obtained). Step (3) is the coordination procedure which consists in adjusting the Lagrange multipliers μ in order to satisfy the constraint $\theta(u) = 0$. μ^k is sent to each subproblem at step (2) in order to compute the dual functional of the problem evaluated at μ^k: $w(\mu^k) = \min_u L(u, \mu^k)$. μ^k is used to find a consensus in the sense of constraint satisfaction. Unfortunately, introducing an augmented Lagrangian with the penalty term $\frac{c}{2}\|\theta(u)\|^2$ would induce the loss of the separability property. Of course, it would be interesting to combine the advantage of augmented Lagrangians (existence of a saddle-point) with separability (leading to parallel or decentralized computations with lower complexity). In Cohen et al. [14], a "norm-linearization technique" has been proposed to meet these two requirements. For that purpose, step (2) of the Uzawa algorithm is simply replaced by

$$u^{k+1} = arg\min_{u} J(u) + \frac{1}{2\epsilon}\|u - u^k\|^2 + < \mu^k + c\theta(u^k), \theta'(u^k).u >, \qquad (12)$$

where $0 < \epsilon < 1/c\tau^2$, with τ being the Lipschitz constant of $\theta(u)$, and $\theta'(u)$ is the Jacobian matrix of θ, while, in the same time, ρ^k can be chosen constant, $\rho^k = \rho$, with $0 < \rho < 2c$.

3.2 Application to state observer design

In our case, in order to circumvent the loss of separability induced by $\frac{c}{2}\|\tilde{s}_i(t) - \sum_{j=1}^{N} F_{ij}\tilde{x}_j(t)\|^2$ and according to

[1] steps ρ^k with $\rho^k \geq 0$, $\sum_{i=0}^{+\infty} \rho^k = +\infty$ and $\sum_{i=0}^{+\infty} (\rho^k)^2 < +\infty$.

(12), we consider the modified Lagrangian

$$\bar{L}_c(\tilde{v},\tilde{s},\tilde{\mu}) = \frac{1}{2}\sum_{i=1}^{N}\int_0^T \{\|C_i\tilde{x}_i(t)-\tilde{y}_i(t)\|^2_{R_i^{-1}} + \|\tilde{v}_i(t)\|^2_{Q_i^{-1}}\}dt + \frac{1}{2}\sum_{i=1}^{N}\|\tilde{x}_i(T)-\underline{x}_i\|^2_{M_i^{-1}}$$

$$+\frac{1}{2\epsilon}\sum_{i=1}^{N}\int_0^T \|\tilde{x}_i(t)-\tilde{x}_i^k(t)\|^2dt + \frac{1}{2\epsilon}\sum_{i=1}^{N}\int_0^T \|\tilde{s}_i(t)-\tilde{s}_i^k(t)\|^2dt + \sum_{i=1}^{N}\int_0^T < \tilde{\mu}_i(t) + c[\tilde{s}_i^k(t) - \sum_{j=1}^{N}F_{ij}\tilde{x}_j^k(t)], \tilde{s}_i(t) - \sum_{j=1}^{N}F_{ij}\tilde{x}_j(t) > dt \quad (13)$$

s.t.

$$\dot{\tilde{x}}_i(t) = -A_i\tilde{x}_i(t) - B_i\tilde{u}_i(t) - E_i\tilde{v}_i(t) - D_i\tilde{s}_i(t)$$

where $0 < \epsilon < 1/c\tau^2$, with τ being the Lipschitz constant of the vector of constraints $\tilde{s}_i(t) - \sum_{j=1}^{N}F_{ij}\tilde{x}_j(t)$, $i = 1, ..., N$. $\bar{L}_c$ is obviously the sum of N independant optimal LQ tracking problems, defined at each dual iteration k, and whose solutions are given by the following Lemma.

Lemma 3.1 *The optimal solution to each LQ tracking problem i, $i = 1, ..., N$*

$$\min_{\tilde{v}_i,\tilde{s}_i,t\in I'} \frac{1}{2}\int_0^T \{\|C_i\tilde{x}_i(t)-\tilde{y}_i(t)\|^2_{R_i^{-1}} + \|\tilde{v}_i(t)\|^2_{Q_i^{-1}}\}dt + \frac{1}{2}\|\tilde{x}_i(T)-\underline{x}_i\|^2_{M_i^{-1}} + \frac{1}{2\epsilon}\int_0^T \|\tilde{x}_i(t)-\tilde{x}_i^k(t)\|^2dt + \frac{1}{2\epsilon}\int_0^T \|\tilde{s}_i(t)-\tilde{s}_i^k(t)\|^2dt + \int_0^T < \tilde{\mu}_i^k(t) + c[\tilde{s}_i^k(t) - \sum_{j=1}^{N}F_{ij}\tilde{x}_j^k(t)], \tilde{s}_i(t) - \sum_{j=1}^{N}F_{ij}\tilde{x}_j(t) > dt \quad (14)$$

$$s.t. \quad \dot{\tilde{x}}_i(t) = -A_i\tilde{x}_i(t) - B_i\tilde{u}_i(t) - E_i\tilde{v}_i(t) - D_i\tilde{s}_i(t)$$

is given, for all $t \in I'$, by

$$\begin{aligned}\tilde{v}_i(t) &= Q_iE_i^T(\tilde{P}_i(t)\tilde{x}_i(t) + \tilde{g}_i(t)),\\ \tilde{s}_i(t) &= s_i^k(t) + \epsilon(D_i^T(\tilde{P}_i(t)\tilde{x}_i(t) + \tilde{g}_i(t)) - \tilde{\mu}_i^k(t)\\ &\quad -c[\tilde{s}_i^k(t) - \sum_{j=1}^{N}F_{ij}\tilde{x}_j^k(t)]),\end{aligned} \quad (15)$$

where $\tilde{P}_i(t)$, $\tilde{g}_i(t)$ and $\tilde{x}_i(t)$ are solutions of the canonical equations

$$\begin{aligned}-\dot{\tilde{P}}_i(t) &= -A_i^T\tilde{P}_i(t) - \tilde{P}_i(t)A_i - \tilde{P}_i(t)E_iQ_iE_i^T\tilde{P}_i(t)\\ &\quad -\epsilon\tilde{P}_i(t)D_iD_i^T\tilde{P}_i(t) + C_i^TR_i^{-1}C_i + \frac{1}{\epsilon}I_d, \quad (16)\\ \tilde{P}_i(T) &= M_i^{-1},\\ -\dot{\tilde{g}}_i(t) &= -A_i^T\tilde{g}_i(t) - \epsilon\tilde{P}_i(t)\tilde{g}_i(t) - C_i^TR_i^{-1}y_i(t)\\ &\quad -\tilde{P}_i(t)D_i(\tilde{s}_i^k(t) - \epsilon(\tilde{\mu}_i^k(t)\\ &\quad +c[\tilde{s}_i^k(t) - \sum_{j=1}^{N}F_{ij}\tilde{x}_j^k(t)]))\\ &\quad -\sum_{j=1}^{N}F_{ji}^T(\tilde{\mu}_j^k(t) + c[\tilde{s}_j^k(t) - \sum_{i=1}^{N}F_{ji}\tilde{x}_i^k(t)])\\ &\quad -\tilde{P}_i(t)B_i\tilde{u}_i(t) - \tilde{P}_i(t)E_iQ_iE_i^T\tilde{g}_i(t)\\ &\quad -\frac{1}{\epsilon}\tilde{x}_i^k(t), \quad (17)\\ \tilde{g}_i(T) &= -M_i^{-1}\underline{x}_i,\\ \dot{\tilde{x}}_i(t) &= -A_i\tilde{x}_i(t) - B_i\tilde{u}_i(t) - E_i\tilde{v}_i(t) - D_i\tilde{s}_i(t), \quad (18)\\ \tilde{x}_i(0) &= -\tilde{P}_i(0)^{-1}\tilde{g}_i(0). \quad (19)\end{aligned}$$

Proof of Lemma 3.1. This optimal LQ tracking problem is well documented in the literature. Equations (16)-(17)-(18)-(19) are easily obtained in closed-loop form both by application of the Pontryagin maximum principle and by considering that the adjoint state $\tilde{p}_i(t)$ has a separable form in both time and space $\tilde{p}_i(t) = \tilde{P}_i(t)\tilde{x}_i(t) + \tilde{g}_i(t)$, where $\tilde{P}_i(t)$ is a symmetric matrix and $\tilde{g}_i(t)$ is a vector of the same dimensions as those of state $\tilde{x}_i(t)$ (see e.g. [19]). Furthermore (19) reflects the fact that each $\tilde{x}_i(0)$ has to minimize $\bar{L}_c(\tilde{v},\tilde{s},\tilde{\mu}^k)$.

Theorem 3.2 *The optimal trajectory solution $\hat{x}_i(t)$, $t \in I$, to each LQ tracking problem (14), $i = 1, ..., N$, expressed in positive time, at each dual iteration k, is given by*

$$\begin{aligned}\dot{W}_i(t) &= W_i(t)A_i^T + A_iW_i(t) + E_iQ_iE_i^T\\ &\quad + \epsilon D_iD_i^T - W_i(t)(C_i^TR_i^{-1}C_i + \frac{1}{\epsilon}I_d)W_i(t), \quad (20)\\ W_i(0) &= M_i.\\ \dot{\hat{x}}_i(t) &= (A_i - W_i(t)C_i^TR_i^{-1}C_i)\hat{x}_i(t)\\ &\quad + W_i(t)C_i^TR_i^{-1}y_i(t) + B_iu_i(t)\end{aligned}$$

$$+ D_i s_i(t) + \frac{1}{\epsilon} W_i(t)(\hat{x}_i^k(t) - \hat{x}_i(t)),$$
$$+ W_i(t)(\sum_{j=1}^{N} F_{ji}^T(\mu_j^k(t) + c[s_j^k(t) - \sum_{i=1}^{N} F_{ji}\hat{x}_i^k(t)])),$$
$$\hat{x}_i(0) = \underline{x}_i, \tag{21}$$

with

$$s_i(t) = s_i^k(t) - \epsilon(\mu_i^k(t) + c[s_i^k(t) - \sum_{j=1}^{N} F_{ij}\hat{x}_j^k(t)]). \tag{22}$$

Proof of Theorem 3.2. Coming back to positive time and using the notations $\bar{x}_i(t) = \tilde{x}_i(T-t)$, $P_i(t) = \tilde{P}_i(T-t)$, $g_i(t) = \tilde{g}_i(T-t)$, $u_i(t) = \tilde{u}_i(T-t)$, $v_i(t) = \tilde{v}_i(T-t)$, $s_i(t) = \tilde{s}_i(T-t)$ and $\mu_i(t) = \tilde{\mu}_i(T-t)$, differential system (16)-(17) given by Lemma 3.1 becomes

$$\begin{aligned}
\dot{P}_i(t) = & -A_i^T P_i(t) - P_i(t)A_i \\
& -P_i(t)E_iQ_iE_i^T P_i(t) \\
& -\epsilon P_i(t)D_iD_i^T P_i(t) + C_i^T R_i^{-1} C_i + \frac{1}{\epsilon} I_d, \qquad (23) \\
P_i(0) = & M_i^{-1}, \\
\dot{g}_i(t) = & -A_i^T g_i(t) - \epsilon P_i(t)D_iD_i^T g_i(t) - C_i^T R_i^{-1} y_i(t) \\
& -P_i(t)D_i(s_i^k(t) - \epsilon(\mu_i^k(t) \\
& +c[s_i^k(t) - \sum_{j=1}^{N} F_{ij}\bar{x}_j^k(t)])) - P_i(t)B_iu_i(t) \\
& -\sum_{j=1}^{N} F_{ji}^T(\mu_j^k(t) + c[s_j^k(t) - \sum_{i=1}^{N} F_{ji}\bar{x}_i^k(t)]) \\
& -P_i(t)E_iQ_iE_i^T g_i(t) - \frac{1}{\epsilon}\bar{x}_i^k(t), \qquad (24) \\
g_i(0) = & -M_i^{-1}\underline{x}_i.
\end{aligned}$$

If we set $W_i(t) = P_i^{-1}(t)$, differential Riccati equation (23) becomes differential Riccati equation (20). We recognize in equation (20) the differential Riccati equation of the optimal LQ state observer, solution of each LQ tracking problem (14), $i = 1, ...N$. Furthermore, the state observer equation is derived by considering estimated state $\hat{x}_i(t) = -W_i(t)g_i(t)$, $\forall t \in I = [0, T]$, by using (20) and (24) (since according to (19), $\hat{x}_i(T) = -W_i(T)g_i(T)$), which leads to subsystem i observer (21), $i = 1, ..., N$. $\hat{x}_i(0) = \underline{x}_i$, since $g_i(0) = -M_i^{-1}\underline{x}_i$ (that is equivalent to $\tilde{g}_i(T) = 0$ given by Lemma 3.1). Finally setting $\bar{x}_i(t) = \hat{x}_i(t)$, and $\bar{x}_i^k(t) = \hat{x}_i^k(t)$, $\forall t \in I$ leads to (22).

3.3 A decentralized LQ state observer (DSO)

Finally, the decentralized optimal LQ state observer may be described as follows:

(1) Initialization ($k = 0$): Choose $\hat{x}_i^0(t)$, $s_i^0(t)$ and $\mu_i^0(t)$, $\forall t \in [0, T]$, $i = 1, ..., N$.

(2) At iteration k: **For each subsystem i,** $i = 1, ..., N$: Integrate on the interval $[0, T]$, each optimal LQ state observer i, $i = 1, ..., N$ to get estimated trajectory $\hat{x}_i^{k+1}(t)$, by using the solution $W_i(t)$ of Riccati equation (20) (computed off-line if M_i is constant):

$$\begin{aligned}
\dot{\hat{x}}_i^{k+1}(t) = & (A_i - W_i(t)C_i^T R_i^{-1} C_i)\hat{x}_i^{k+1}(t) + B_iu_i(t) \\
& + W_i(t)C_i^T R_i^{-1} y_i(t) + D_i s_i^{k+1}(t) \\
& + W_i(t)(\sum_{j=1}^{N} F_{ji}^T(\mu_j^k(t) + c[s_j^k(t) \\
& \qquad - \sum_{i=1}^{N} F_{ji}\hat{x}_i^k(t)])) \\
& + \frac{1}{\epsilon} W_i(t)(\hat{x}_i^k(t) - \hat{x}_i^{k+1}(t)), \\
\hat{x}_i^{k+1}(0) = & \underline{x}_i. \qquad (25)
\end{aligned}$$

where

$$s_i^{k+1}(t) = s_i^k(t) - \epsilon(\mu_i^k(t) + c[s_i^k(t) - \sum_{j=1}^{N} F_{ij}\hat{x}_j^k(t)]). \tag{26}$$

(3) **Coordination:** Update the μ_i's:

$$\mu_i^{k+1}(t) = \mu_i^k(t) + \rho(s_i^{k+1}(t) - \sum_{j=1}^{N} F_{ij}\hat{x}_j^{k+1}(t)), \tag{27}$$

$\forall t \in [0, T]$, $i = 1, ..., N$.

(4) if $\sum_{i=1}^{N} \|\mu_i^{k+1} - \mu_i^k\|^2 < \epsilon_c^2/\rho^2$ then stop, else $k+1 \to k$ and go to step 2).

The following theorem can be stated:

Theorem 3.3 • **Convergence:** *if $0 < \epsilon < 1/c\tau$ where τ is the Lipschitz constant of the interconnection constraints and $0 < \rho < 2c$, then the sequence $(\hat{x}_i^k, s_i^k, \mu_i^k)$, $i = 1, ..., N$ generated by algorithm (DSO) is bounded and the related cluster point is a saddle-point of L_c.*

• **Qualification of the solution:** *The obtained cluster point is solution to the decentralized optimal LQ state estimation problem (6), which is equivalent to a centralized finite-horizon Kalman filter for system (1)-(2).*

Proof of Theorem 3.3. The first statement is proved by using Theorem 15 in Cohen et al. [14], where the assumptions on convexity of both the cost functionals and the constraints, and the so-called Constraint Qualification Condition are satisfied in our case. The second statement is proved by verifying that the cluster point satis-

fies the necessary conditions for optimality of problem (6) (due to the lack of place, this is left to the reader).

Some Remarks:

- This algorithm is well suited for decentralized state estimation through communication networks, such as sensor networks, since each of the N interconnected agents has to compute its own state estimation on the basis of information $(\hat{x}_j^k(t), \mu_j^k(t))$, $\forall t \in [0,T]$, provided by the others agents j connected to it and then has to send its computation results $(\hat{x}_i^{k+1}(t), \mu_i^{k+1}(t))$, $\forall t \in [0,T]$ at each iteration k, only to these connected agents $j \in C_i$, where C_i is the set of indices of agents connected to i. For n and N large, *the computational complexity can be dramatically reduced from a n-dimensional LQ estimation problem to N n_i-dimensional LQ estimation problem solved in parallel with additional $(n_i + n_{s_i})$-dimensional exchange data $(\hat{x}_i(t), s_i(t), \mu_i(t))$, $\forall t \in [0,T]$, where $n_{[1,9]s_i}$ is vector $s_i(t)$ dimension.*
- *Step* (2) *of algorithm (DSO) is equivalent to the computation of N small-scale finite horizon Kalman filters.*
- Coordination step (3) can be decentralized through each agent i, without needing a central coordinator agent.
- Existence of a bounded and positive definite matrix $P_i(t)$ over the interval $[0,T]$ is ensured by the regularization term $\frac{1}{\epsilon} I_d$ in the Riccati equation (23), *even if pairs (C_i, A_i) are not observable.* This assumption is often needed in other approaches (see [6] for instance).
- The main parameter to be tuned is the regularization coefficient c. It is well known (see [14] for example) that c must be chosen not too small to ensure existence of a saddle-point in the case of non-strongly convex problems, but not too large to avoid bad numerical conditioning in solving step (2) of the price-decomposition-coordination algorithm.
- The practical implementation of this continuous-time decentralized observer may be easily performed by using a piece-wise constant discretization of the interconnection constraints $s_i(t) - \sum_{j=1}^{N} F_{ij} \bar{x}_j(t)$ and the related Lagrange multipliers $\mu_i(t)$, induced by a time discretization of the interval $[0,T]$, while the differential equations (including the Riccati equations) may be solved by using standard integration schemes, together with an approximation of the cost function integral by an appropriate quadrature scheme.

4 A MOVING HORIZON STATE OBSERVER BASED ON ALGORITHM (DSO)

To get on-line state estimation rather than batch estimation, algorithm (DSO), which provides the exact solution to a finite horizon Kalman filter, can be used as the core of a moving horizon observer approach, by considering $\int_{t-T}^{t}$ instead of $\int_{0}^{T}$. The receding horizon observer uses only measurements on the moving horizon $[t-T, t]$ and discards past measurements prior to the current moving horizon (see [17] for instance). Here a sampled moving horizon observer with observation horizon $T = N_s \tau_s$, that updates the estimated state at time samples $t_l = l\tau_s$, where τ_s denotes the measurement sampling period, is defined as follows:

$$\hat{x}_i(t_l) = \mathcal{S}_i(y_{i\,t_{l-N}}^{\,t_l}),\ i = 1, ..., N,\ l > N_s \tag{28}$$

where $\mathcal{S}_i(y_{i\,t_{l-N}}^{\,t_l})$ denotes the estimated state obtained at time t_l from iterative algorithm (DSO) [2], based on the output measurements $y_i(\tau)$, $\tau \in [t_{l-N}, t_l]$, denoted as $y_{i\,t_{l-N}}^{\,t_l}$, for each subsystem i, $i = 1, ..., N$. At time sample t_l, each observer (25) is initialized with estimated state $\hat{x}_i(t_l)$ computed at time sample t_{l-1} (i.e. on moving interval $[t_{l-1-N}, t_{l-1}]$).

4.1 Convergence of the moving horizon state estimation

In this section, we will consider two particular versions of the tuning parameters T and M_i, $i = 1, ..., N$ leading to simple convergence results:

- **Version 1: Deadbeat property:** Following [22], moving horizon observer (28) exhibits a deadbeat property provided that $T > 0$ and $M_i^{-1} = 0$, $i = 1, ..., N$ (assumption of infinite covariance of the initial state). The implementation of this version requires a new filter form for (25) which can be derived by using change of coordinates $\nu_i(t) = W_i^{-1} \hat{x}_i(t)$ (see eqn. (9) in [22] for more details). T may be chosen arbitrarily small since system (1) is time-invariant and the initial estimated state of each observer (25) may be anything at each time sample t_l.
- **Version 2: Asymptotic convergence property:** Choosing $T = \tau_s$ $(N_s = 1)$, $\underline{x}_i = \hat{x}_i(t_{l-1})$, where $\hat{x}_i(t_{l-1})$ denotes the local state estimate computed at previous time sample t_{l-1}, and $M_i = W_i(t_l)_{t_{l-1}}$, $i = 1, ..., N$, where $W_i(t_l)_{t_{l-1}}$ denotes the solution at time sample t_l of Riccati equation (20) computed at previous time sample t_{l-1}, leads to an asymptotic convergence property. In that case, Riccati equations (20) have to be computed on-line, if the estimation horizon T is unknown. However the computational complexity of steps (2) and (3) of algorithm (DSO) is reduced to a minimum since the horizon is limited to one time sample. Since algorithm (DSO) computes the

[2] At each time sample t_l, algorithm (DSO) can be initialized at $k = 0$ with any initial trajectory. However experiments show the use of the optimal solution computed at t_{l-1} can significantly reduce the number of dual iterations, i.e., increase the convergence speed of algorithm (DSO).

exact solution to a Kalman filter at each time sample, as $l \to +\infty$, the moving horizon observer solution tends to the solution of the infinite horizon Kalman filter, which is known for providing state estimation error dynamics stability under the assumption that the overall system is observable.

5 AN ILLUSTRATIVE EXAMPLE

The main objective is to build a state observer in order to estimate the water velocity distribution along an irrigation canal from distributed level measurements. In practice, using a state observer for such systems is very interesting, since it is very expensive to use velocity or flow rate sensors in addition to level sensors. It is assumed that a $N+1$-sensor network has been distributed along the canal stretch $[0, L]$, which consists in $N-1$ sensor-nodes equipped with only a water level sensor, while water flow velocity (or flow rate) are also available at the upstream and downstream nodes via direct measurements or hydraulic laws. The sensor-nodes have some limited computational capacities and are able to communicate with a SCADA master station using for instance an ethernet link, according to Fig. 1.

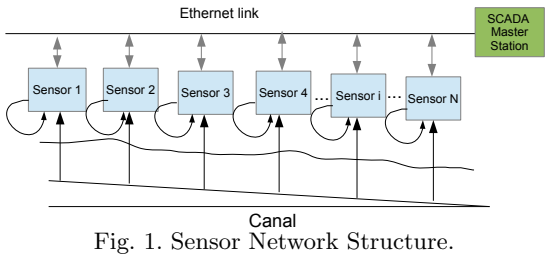

Fig. 1. Sensor Network Structure.

A spatial discretization of Saint Venant equations governing open-channel water flow dynamics, is performed thanks to a finite-difference scheme with spatial sample dx, followed by a linearization of the discretized system around an equilibrium point of Saint Venant equations, which lead to the following observable linear system:

$$\begin{aligned}
\dot{h}_i(t) &= \frac{1}{dx}(v_i(t) - v_{i+1}(t)),\ i = 0, ...N-1,\\
\dot{v}_i(t) &= \frac{1}{dx}(\alpha_{i-1}h_{i-1}(t) + \beta_{i-1}v_{i-1}(t) - \alpha_i h_i(t) - \beta_i v_i(t))\\
&\quad +\delta_i h_i(t) + \gamma_i v_i(t),\ i = 1, ..., N,\\
y_i(t) &= h_i(t),\ i = 1, ..., N-1
\end{aligned} \tag{29}$$

where $h_i(t)$, and $v_i(t)$ stand for the deviation values from the equilibrium of the water depth and velocities, expressed at each spatial coordinate x_i. More details on this derivation are available in [16]. A natural decomposition of this system under form (3)-(5) can be obtained for each sensor i, $i = 1, ..., N$, as follows:

$$\begin{aligned}
\dot{h}_i(t) &= \frac{1}{dx}(v_i(t) - s_i^1(t)),\\
\dot{v}_i(t) &= \frac{1}{dx}(\alpha_{i-1}s_i^2(t) + \beta_{i-1}s_i^3(t) - \alpha_i h_i(t) - \beta_i v_i(t))\\
&\quad +\delta_i h_i(t) + \gamma_i v_i(t),\\
y_i(t) &= h_i(t),\ i = 1, ..., N-1,\\
s_i(t) &= (s_i^1(t), s_i^2(t), s_i^3(t))^T = (v_{i+1}(t), h_{i-1}(t), v_{i-1}(t))^T\\
\dot{h}_0(t) &= \frac{1}{dx}(v_0(t) - s_1^2(t)),\\
\dot{v}_0(t) &= u_1(t),\\
y_0(t) &= (h_0(t), v_0(t))^T,\\
s_0(t) &= s_0^1(t) = v_1(t),\\
\dot{h}_N(t) &= u_2(t),\\
\dot{v}_N(t) &= \frac{1}{dx}(\alpha_{N-1}s_N^1(t) + \beta_{N-1}s_N^2(t)\\
&\quad -\alpha_N h_N(t) - \beta_N v_N(t)) + \delta_N h_N(t) + \gamma_N v_N(t),\\
y_N(t) &= (h_N(t), v_N(t))^T,\\
s_N(t) &= (s_N^1(t), s_N^2(t))^T = (h_{N-1}(t), v_{N-1}(t))^T
\end{aligned}$$

The interconnection variables correspond to states of the nearest neighbours of i only. Consequently, *the data exchanges are limited to the nearest neighbors of each sensor.* Table 1 summarizes the parameters of the test system, which corresponds to a 10 km-long canal. The decentralized observer parameters used in the simulations are given by table 2, where I_n is the $(n \times n)$ identity matrix and in addition τ is the Lipschitz constant of the constraints, $\rho = 1.99c$ and $\epsilon = \frac{0.99}{c\tau^2}$.

Table 1
Test system parameters.

N	dx	h_e	u_e
20	$500m$	$h_e(x) = 1m$	$u_e(x) = 0.839m/s$

Table 2
Decentralized observer parameters.

Q_i	R_i	M_i	c	ϵ_c
$10^{-1}I_2$	10^{-6} or $10^{-6}I_2$	$10^{-1}I_2$	10^{-4}	$10^{-7}3N \times N_s$

In this cas study, version 2 of the (DSO)-based decentralized moving-horizon observer (DMHO1), (where $N_s = 1$), and a version with a larger moving horizon $N_s = 4$ (DMHO2) were implemented with sampling time $\tau_s = 480s$ on an estimation horizon of $M = 20$ time samples. The simulations were performed on a desktop with Intel Core i5-2320 @ 3.00GHz and 6 Go RAM under SCILAB software [3]. Fig. 2 and 3 provide a comparison between the simulated linearized system and the two versions of the observer, with a Gaussian noise of zero mean and a standard deviation of 10^{-3} added to the measured outputs. Water velocities are effectively estimated, after a

[3] SCILAB is an open source software for numerical computation, trademark of Scilab Entreprises SAS 2013.

10% increase of the upstream water velocity. For both versions of the observer, the average number of dual iterations at each sample time is equal to 57.

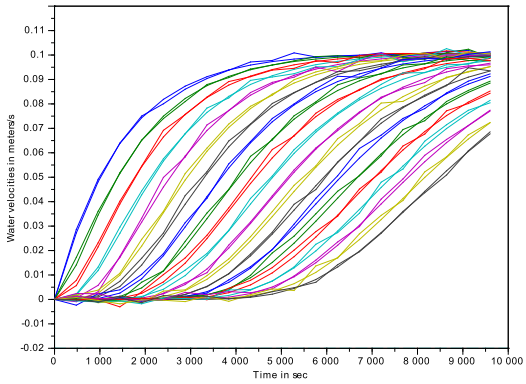

Fig. 2. Comparison between the true water velocities and the estimated ones - with DMHO1.

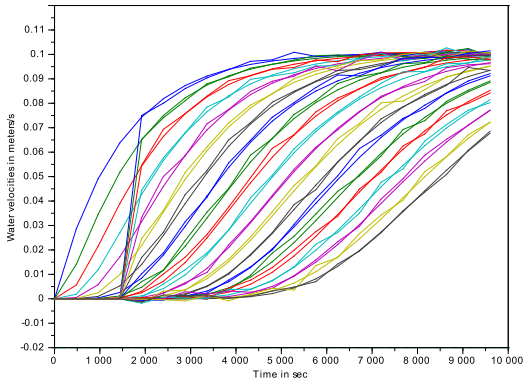

Fig. 3. Comparison between the true water velocities and the estimated ones - with DMHO2.

5.1 Comparisons with some alternative methods

Using the same case study with the same covariance matrices as defined in table 2, the here-proposed approaches (DMHO1 and DMHO2) are successively compared with the *Centralized Infinite-Horizon Kalman Filter* (CIHKF), the *Centralized Information Filter* (CIF) described in [4] (page 4922). The centralized infinite-horizon Kalman filter provides the optimal LQ state estimation in the sense of the minimal variance of the state estimation error in the presence of both state and output noises defined by the covariance matrices in table 2 or, equivalently, the solution of LQ problem (6) when $T \to +\infty$. The Centralized Information Filter is a partially distributed implementation of the finite-horizon Kalman filter: At each sensor level, the so-called local observation variables are computed and sent to a central location which computes updates of both the global state estimate and the Riccati equation covariance matrix at each time sample. When the number of sensor nodes N is large, the number of communications and the $O(n^3)$ computations at the central location (where n is the state number) are some serious limitations of the approach, together with large latency and single point of failure disadvantages. In comparison, approach (DMHO) does not require a central location, and all the $O(n^3)$ computations are reduced to $\sum_{i=1}^{N} O(n_i^3)$ where n_i is the state dimension of subsystem i, $i = 1, ..., N$. Communications are also limited to connected neighbors of each subsystem. Table 3 provides some comparisons of the root-mean-square error of estimation ($RMSE$: $\frac{1}{M}\sum_{i=1}^{M} \|x(k) - \hat{x}(k)\|^2$), the average time required to compute the overall state estimation (denoted as T_{node}) by each sensor node (subsystem). In the case of CIF, the time needed to compute the Riccati covariance matrices over the estimation horizon (denoted as T_c) by the central location is also provided.

Table 3
Comparisons (with $N = 20$, $\tau = 480s$, $M = 20$, and the parameters in Tables 1 and 2).

Approach	$RMSE$	T_{node} (in sec)	T_c (in sec)
DMHO1	6.16×10^{-5}	210	n.a.
DMHO2	6.13×10^{-5}	541	n.a.
CIHKF	4.44×10^{-5}	n.a.	n.a.
CIF	6.4×10^{-5}	78	29549

All the approaches exhibited similar RMSE, with a better estimation provided by CIHKF. CIF is the far less effective approach in terms of computation time needed at the central location. For both DMHO1 and DMHO2, real-time implementation is possible since each node can compute the current estimate in less than $\tau = 480s$ (the worst average computation time per node is given by $541/M = 27.05s$).

6 CONCLUSIONS

A decentralized optimal LQ state observer design suitable for a sensor-network-based state estimation of large-scale interconnected systems has been proposed. This observer design relies on both the use of an augmented Lagrangian formulation and a price-decomposition-coordination algorithm. Under mild assumptions, the algorithm converges to the solution of the centralized optimal LQ observer (finite-horizon Kalman filter). This algorithm is the core of a decentralized moving horizon state estimator. A 20 sensor-network based hydraulic canal estimation problem illustrated the effectiveness of the proposed approach, which was also

compared to existing approaches. A very similar version may be derived in the framework of discrete-time linear interconnected systems. The extension to the case of nonlinearly-interconnected linear systems is also possible, with very similar communication and computation requirements. Further researches will be devoted to the analysis of the impact of communication noises or losses on the observer performance.

References

[1] D. Looze, P. Houpt, N. Sandell, and M. Athans (1978). On decentralized estimation and control with application to freeway ramp metering. In *IEEE Transactions on Automatic Control, Vol. AC-23, 2* , 268-275.

[2] C. W. Sanders, E. C. Tacker and T. D. Linton (1978). Specific structures for large scale state estimation algorithms having information exchange. In *IEEE Transaction on Automatic Control, Vol. 23, pp. 255?260.*

[3] D. Siljak (1991) Decentralized Control of Complex Systems. In *Academic Press, 1991.*

[4] U. Khan, and J. Moura (2008). Distributing the Kalman filter for large-scale systems. In *IEEE Transactions on Signal Processing, Vol. 56, 10* , 4919-4935.

[5] S. Stankovic, M. Stankovic, and D. Stipanovic (2009). Consensus based overlapping decentralized estimator. In *IEEE Transactions on Automatic Control, Vol. 54, 2* , 410-415.

[6] M. Farina, G. Farrari-Trecate, and R. Scattolini (2010). Moving-horizon partition-based state estimation of large scale systems. In *Automatica, vol. 46*, 910–18.

[7] A. Hess and A. Rantzer (2010). Distributed Kalman filter algorithms for self-localization of mobile devices. In *Proceedings of 13th ACM International Conf. on Hybrid Systems: Computation and Control*, Stockholm, Sweden.

[8] J. Maestre, P. Giselsson, and A. Rantzer (2010). Distributed receding horizon Kalman filter. In *Proceedings of 49th IEEE Conf. on Decision and Control*, Atlanta, GA, USA.

[9] M. Mansouri, L. Boutat-Baddas, M. Darouach, and H. Messaoud (2010). Decentralized observers design for descriptor systems. In *Proceedings of the 18th Mediterranean Conference on Control And Automation*, Marrakech, Morocco.

[10] Takayuki Ishizaki, Yukihiro Sakai, Kenji Kashima, and Junichi Imura (2011). Hierarchical decentralized observer design for linearly coupled network systems. In *Proceedings of 50th IEEE CDC-ECC*, Orlando,FL, USA.

[11] K. Martensson and A. Rantzer (2011). A scalable modularized synthesis method for distributed Kalman filters. In *Proceedings of IFAC World Congress*, Milano, Italy.

[12] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci (2002). Wireless sensor networks: A survey. In *Computer Networks 38*, 393–422.

[13] M. Ghanem, Y. Guo, J. Hassard, M. Osmond, M. Richards (2004). Sensor Grids For Air Pollution Monitoring. In *Proc. 3rd UK e-Science All Hands Meeting.*

[14] G. Cohen and D.L. Zhu (1984). Decomposition coordination methods in large scale optimization problems. The nondifferentiable case and the use of augmented Lagrangians. In *J.B. Cruz (Ed.), Advances in Large Scale Systems, Vol. I*, 203–266, JAI Press, Greenwich, Connecticut.

[15] D. Georges (2006). Distributed model predicted control via decomposition-coordination techniques and the use of an augmented lagrangian. In *IFAC Workshop on NMPC for Fast Systems*, 111–116, Grenoble, France.

[16] D. Georges (2013). State observer design for open-channel hydraulic systems based on a sensor network. In *Internal Report*, GIPSA-lab, Grenoble Institute of Technology, France.

[17] M. Alamir (2007). Nonlinear Moving Horizon Observers: Theory & Real-Time Implementation. In *Nonlinear Observers and Applications, Gildas Besancon (Ed). Lecture Notes on Communication and Information Science.* Springer-Verlag-Series.

[18] G. Bornard, F. Celle-Couenne and G. Gilles (1995). Observability and observers. In *Nonlinear Systems - T.1, Modeling and Estimation*, Chapman & Hall, London.

[19] B.D.O. Anderson and J.B. Moore (1990). Optimal control. Linear Quadratic Methods. In *Prentice Hall, Englewood Cliffs, New Jersey.*

[20] H. Kwakernaak and R. Sivan (1972). Linear optimal control systems. In *Wiley-Interscience, New York.*

[21] K. Arrow, L. Hurwicz and H. Uzawa (1958). Studies in Linear and Nonlinear Programming. In *Stanford University Press*, Stanford.

[22] W. H. Kwon, P. S. Soo and P. Park (1999). A receding horizon Kalman FIR filter for linear continuous-time systems. In *IEEE Transactions on Automatic Control, Vol. 44, 11* , 2115-2120.