2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

A Consensus Based Overlapping Decentralized Estimator in Lossy Networks: Stability and Denoising Effects

Srdjan S. Stanković, Miloš S. Stanković and Dušan M. Stipanović

Abstract— **In this paper a new algorithm for discrete-time overlapping decentralized state estimation has been analyzed. It is based on a combination of Kalman filters implemented by local agents using intermittent observations and a dynamic consensus strategy connecting the agents with possible communication faults. Under general conditions concerning the agent resources and the network topology, sufficient conditions for the asymptotic stability in the sense of bounded mean-square estimation error are derived. It is also demonstrated how the complexity of the multi-agent network contributes to the suppression of the measurement noise influence.**

I. INTRODUCTION

Recently, much research has been done in the field of *decentralized state estimation* of complex dynamic systems. When dealing with these problems, different structures have been considered, including totally decentralized, partially decentralized and hierarchical structures. In all these approaches a large scale system is modelled as an interconnection of *subsystems* where each subsystem has a decision maker (intelligent *agent*) associated with it. Depending on the available resources, each agent might have different sensor characteristics, different models of the system and its environment, and each agent might communicate with different sets of other agents. The main principles and structures for decentralized estimation can be found in *e.g.* [1], [2], [3], [4], [5]. However, none of the existing methodologies is able to provide a systematic and general method for designing a scheme for inter-agent communication without having a strong fusion center. Also, the important practical problems of intermittent observations and communication faults have not been treated in this context.

Some important results in the area of distributed asynchronous iterations in parallel computation and distributed optimization were already obtained in the 80's (*e.g.* [6], [7], [8]). On the other hand, much research has been done recently in the field of multi-agent systems related to sensor networks, with numerous applications. One approach to solving a large class of problems is based on *agreement* or *consensus* strategies between the agents (see, *e.g.* [9], [10], [11], [12], [13], [14], [15]). The consensus problem can be related to the decentralized state estimation problem either implicitly, through the very definition of the consensus algorithms (*e.g.*, see [15]), or explicitly, where the dynamic consensus strategy between multiple agents is used for obtaining, on the basis of averaging, optimal parameter or state estimates (*e.g.*, see [16], [17], [18]). However, none of these approaches provide a general cooperation scheme between the local estimators for the overlapping decentralized estimation problem.

This paper is a continuation of [19], in which a novel state estimation algorithm for complex linear discrete-time systems is proposed. It is based on: (1) decomposition of the large system into overlapping subsystems and implementation of local state estimators, for each subsystem, by intelligent agents according to their sensing and computing resources; (2) application of a consensus strategy which provides the global state estimates to all the agents in the network; (3) taking into account influence of the possible intermittent observations and communication errors. The paper is organized as follows. The definition of the problem is given in Section 2. In Section 3, the proposed estimation algorithm is described. The algorithm can be considered as a discrete-time version of the state estimation algorithm proposed in [20], or an extension of the parameter estimation algorithm proposed in [21]. It is structurally similar to the distributed computation algorithm proposed in [7] and [8]. In Section 4 the stability analysis of the proposed scheme is presented. Sections 2, 3 and 4 are technically the same as in [19]; they are repeated here for completeness of the presentation. Section 5 provides an insight into an interesting and important problem of the influence of the network topology to denoising, that is, to the measurement noise suppression. It is shown that the network complexity is crucial in determining denoising capabilities of the algorithm.

II. OVERLAPPING DECENTRALIZED ESTIMATION

Let a discrete-time stochastic system be represented by

$$\mathbf{S}: \quad \begin{aligned} x(t+1) &= Fx(t) + Ge(t) \\ y(t) &= Hx(t) + v(t), \end{aligned} \tag{1}$$

where t is the discrete-time instant, $x = (x_1, \ldots, x_n)^T$, $y = (y_1, \ldots, y_p)^T$, $e = (e_1, \ldots, e_m)^T$ and $v = (v_1, \ldots, v_p)^T$ are its state, output, input and measurement noise vectors, respectively, while F, G and H are constant $n \times n$, $n \times m$ and $p \times n$ matrices, respectively. It is assumed that $\{e(t)\}$ and $\{v(t)\}$ are white zero-mean sequences of independent vector random variables with covariance matrices Q and R.

Our task is to consider the problem of *decentralized estimation* of the state x of $\mathbf{S}$. We shall assume that N *autonomous agents* generate their estimates of the state of

S. S. Stanković who is the corresponding author is with the Faculty of Electrical Engineering, University of Belgrade, 11000 Belgrade, Serbia, Phone: 381-11-337-0150, Fax: 381-11-324-8681; E-mail: stankovic@etf.bg.ac.yu

M. S. Stanković and D. M. Stipanović are with the Department of Industrial and Enterprise Systems Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA, Phone: 1-217-244-0907, Fax: 1-217-244-5705; E-mail: mstanko3@uiuc.edu, dusan@uiuc.edu

$\mathbf{S}$ using locally available measurements, specific *a priori* knowledge about the system model or its parts, and estimates of other agents communicated in real time.

Formally speaking, the i-th agent measures the p_i-dimensional vector $y^{(i)} = (y_{l_1^i}, \dots, y_{l_{p_i}^i})^T$, composed of the set of components of y with indices contained in the agent's *output index set* $I_i^y = \{l_1^i, \dots, l_{p_i}^i\}$, $l_1^i, \dots, l_{p_i}^i \in \{1, \dots p\}$, $l_1^i < \dots < l_{p_i}^i$, $p_i \leq p$. We shall assume that vector $y^{(i)}$ satisfies the following relation

$$y^{(i)}(t) = H^{(i)} x^{(i)}(t) + v^{(i)}(t), \tag{2}$$

where $x^{(i)}$ is an n_i-dimensional vector composed of the components of x selected by the agent's *state index set* $I_i^x = \{j_1^i, \dots, j_{n_i}^i\}$, $j_1^i, \dots, j_{n_i}^i \in \{1, \dots n\}$, $j_1^i < \dots < j_{n_i}^i$, $n_i \leq n$, $C^{(i)}$ is a constant $p_i \times n_i$ matrix and $v^{(i)}$ a measurement noise vector containing the components of v selected by I_i^y, having covariance matrix $R^{(i)}$ (which can be easily derived from R). According to the definition of $x^{(i)}$, we introduce an $n_i \times n_i$ matrix $F^{(i)}$ containing those components of F that are selected by the pairs of indices defined by $I_i^x \times I_i^x$. In an analogous way we can obtain matrix $G^{(i)}$, composed of n_i rows of matrix G selected by I_i^x. Consequently, the *local models* of $\mathbf{S}$ utilized by the agents are defined by

$$\mathbf{S_i}: \quad \begin{aligned} x^{(i)}(t+1) &= F^{(i)} x^{(i)}(t) + G^{(i)} e(t), \\ y^{(i)}(t) &= H^{(i)} x^{(i)}(t) + v^{(i)}(t), \end{aligned} \tag{3}$$

for $i = 1, \dots, N$, where dynamic systems $\mathbf{S_i}$ represent *overlapping subsystems* of $\mathbf{S}$ ([3], [22]).

Starting from the model $\mathbf{S_i}$ and the accessible measurements $y^{(i)}$, each agent is assumed to be able to generate its own local estimate $\hat{x}^{(i)}$ of the vector $x^{(i)}$ using an estimator designed on the basis of the local model (3). Having in mind the stochastic nature of the system model $\mathbf{S}$, the following local steady-state Kalman filters will be assumed to be implementable by each agent [23]:

$$\mathbf{\bar{E}_i}: \quad \hat{x}^{(i)}(t+1|t) = F^{(i)} \hat{x}^{(i)}(t|t-1) + \gamma_i(t) F^{(i)} L^{(i)} [y^{(i)}(t) - H^{(i)} \hat{x}^{(i)}(t|t-1)], \tag{4}$$

where $L^{(i)}$ is the steady state Kalman gain, while $\gamma_i(t)$ is a scalar equal to 1 when the i-th agent receives measurements $y^{(i)}$, and 0 otherwise. We shall assume that the pairs $(F^{(i)}, G^{(i)} Q^{\frac{1}{2}})$ are completely stabilizable and the pairs $(F^{(i)}, H^{(i)})$ completely detectable, so that the the state matrices $F^{(i)} - L^{(i)} H^{(i)}$ of the estimators (4) are asymptotically stable, $i = 1, \dots, N$ [23], [24].

The above described overlapping decentralized estimators defined by (4) generate a set of *overlapping estimates* $\hat{x}^{(i)}$. If the final goal is to enable each agent to get an estimate $\hat{x}$ of the whole state vector x of $\mathbf{S}$, additional strategies have to be combined with the local estimators (*e.g.*, see [1], [2], [4], [5], [3]). However, it is typical that such approaches require a kind of centralized strategy or special, model dependent communications.

III. CONSENSUS-BASED ESTIMATOR

Our goal is to design an algorithm which provides to all the agents in the network reliable estimates of all the components of the whole state vector x using locally generated estimates $\hat{x}^{(i)}$ and a decentralized communication strategy, in spite of missing measurements and communication faults. We propose the following algorithm based on the synergy of local estimators of Kalman filtering type and a *consensus scheme* defining communications between the agents:

$$\mathbf{E_i}: \begin{aligned} \xi_i(t|t) &= \xi_i(t|t-1) + \gamma_i(t) L_i [y^{(i)} - H_i \xi_i(t|t-1)] \\ \xi_i(t+1|t) &= \textstyle\sum_{j=1}^N C_{ij}(t) F_j \xi_j(t|t) \end{aligned} \tag{5}$$

for $i = 1, \dots, N$, where ξ^i is an estimate of x generated by the i-th agent. F_i is an $n \times n$ matrix with at most $n_i \times n_i$ nonzero elements that are equal to those of $F^{(i)}$, but are placed at the indices defined by $I_i^x \times I_i^x$, L_i is an $n \times p_i$ matrix obtained similarly as F_i, in such a way that its nonzero elements are those of $L^{(i)}$ placed row by row at row-indices defined by I_i^x, H_i is a $p_i \times n$ matrix composed of the entire rows of H selected by I_i^y (notice that we have $H^{(i)} x^{(i)} = H_i x$). We shall assume that $C_{ij}(t)$ are $n \times n$ time-varying gain matrices of the consensus scheme given in the form $C_{ij}(t) = k_{ij}(t) K_{ij}$, where $k_{ij}(t) = 1$ when the directed communication from the node j to the node i exists, and $k_{ij}(t) = 0$ otherwise; K_{ij} are diagonal matrices with nonnegative elements, giving weights to the estimates $F_j \xi_j(t|t)$ that the agents send to each other (for example, its elements can be inversely proportional to the variances of the local estimates). Formally, we shall assume that $\{k_{ij}(t)\}$, $i, j = 1, \dots, N, i \neq j$, are mutually independent scalar sequences of independent binary random variables, satisfying $P\{k_{ij}(t) = 1\} = p_{ij}$ and $P\{k_{ij}(t) = 0\} = 1 - p_{ij}$ for $i \neq j$, and that that $k_{ii}(t) = 1$, $i = 1, \dots, N$; the $N \times N$ matrix $K(t) = [k_{ij}(t)]$ determines connections between the agents at time t. Also, we shall assume that $\{\gamma_i(t)\}$ is a sequence of independent binary random variables independent of $\{k_{ij}(t)\}$, $i, j = 1, \dots, N, i \neq j$, such that $P\{\gamma_i(t) = 1\} = p_{ii}$ and $P\{\gamma_i(t) = 0\} = 1 - p_{ii}$. Furthermore, we shall introduce a random vector Ξ_t composed of N^2 binary components: $N(N-1)$ elements $k_{ij}(t)$ ($i, j = 1, \dots, N, i \neq j$) and N elements $\gamma_i(t)$. This vector is, by assumption, generated on the basis of Bernoulli trials, *i.e.*, $\{\Xi_t\}$ represents a sequence of independent random vectors. Let π_r, $i = 1, \dots, \nu = 2^{N^2}$, be the probabilities of all possible realizations $\Xi^{[r]}$ of Ξ_t, , $r = 1, \dots, \nu$. Assume that the $nN \times nN$ "consensus matrix" $\tilde{C}(t) = [C_{ij}(t)]$, $i, j = 1, \dots, N$ is *row-stochastic* for all t. Introduce also $\Phi_i(t) = F_i - \gamma_i(t) L_i H_i$, $\tilde{F}_E = \mathrm{diag}\{F_1, \dots, F_N\}$ and $\tilde{\Phi}(t) = \mathrm{diag}\{\Phi_1(t), \dots, \Phi_N(t)\}$, as well as $\tilde{A}(t) = \tilde{C}(t)\tilde{\Phi}(t)$. Denote by $\tilde{A}_{[r]}$, $\tilde{C}_{[r]}$ and $\tilde{\Phi}_{[r]}$ realizations of $\tilde{A}(t)$, $\tilde{C}(t)$ and $\tilde{\Phi}(t)$ resulting from different realizations $X^{[r]}$ of X_t, $r = 1, \dots, \nu$.

It can be observed that the algorithm represents a combination of: a) *decentralized overlapping estimators* represented by (4), and b) *the consensus scheme* tending to make the local estimates ξ^i as close as possible (*e.g.*, see [8], [7], [9],

[10], [11], [12], [13], [14], [15]). The algorithm reduces to the local estimators when the "consensus part" is eliminated ($K_{ij} = 0$, $i \neq j$). The "consensus part" alone asymptotically provides $\xi^i = \xi$ under a proper choice of the matrices K_{ij}, where ξ is a weighted sum of the *a priori* estimates $\xi(t_0)$, and t_0 the initial time instant ([14], [15]). Notice that the estimator reminds structurally of the discrete time distributed optimization algorithm proposed in [8], [7], [6]. Also, it can be considered as a discrete time counterpart of the continuous time overlapping decentralized estimator proposed in [20]. Additionally, it represents a generalization of the parameter estimator based on stochastic approximation proposed in [21].

By defining $\hat{X}(t|t) = \mathrm{vec}\{\xi_1(t|t), \ldots, \xi_N(t|t)\}$ and $\hat{X}(t+1|t) = \mathrm{vec}\{\xi_1(t+1|t), \ldots, \xi_N(t+1|))\}$, we can obtain a compact formulation of the proposed algorithm

$$\begin{aligned}\hat{X}(t|t) &= \hat{X}(t|t-1) + \tilde{\Gamma}(t)\tilde{L}[Y(t) - \tilde{H}\hat{X}(t|t-1)] \\ \hat{X}(t+1|t) &= \tilde{C}(t)\tilde{F}_E\hat{X}(t|t),\end{aligned} \quad (6)$$

where $Y(t) = \mathrm{vec}\{y^{(1)}(t), \ldots, y^{(N)}(t)\}$, $\tilde{L} = \mathrm{diag}\{L_1, \ldots, L_N\}$, $\tilde{\Gamma}(t) = \mathrm{diag}\{\gamma_1(t), \ldots, \gamma_N(t)\}$ and $\tilde{H} = \mathrm{diag}\{H_1, \ldots, H_N\}$. Further, for the prediction error $\varepsilon(t+1|t) = \hat{X}(t+1|t) - X(t+1)$, we obtain

$$\begin{aligned}\varepsilon(t+1|t) &= \tilde{A}(t)\varepsilon(t|t-1) + \tilde{C}(t)(\tilde{F}_E - \tilde{F}) + \\ &+ \tilde{C}(t)\tilde{\Gamma}(t)\tilde{L}\tilde{H}V(t) - E(t),\end{aligned} \quad (7)$$

where $\tilde{F} = \mathrm{diag}\{F, \ldots, F\}$, $V(t) = \mathrm{vec}\{v^{(1)}(t), \ldots, v^{(N)}(t)\}$ and $E(t) = \mathrm{vec}\{e(t), \ldots, e(t)\}$. Finally, we have the following state space system-estimator model:

$$\begin{aligned}Z(t+1) &= \left[\begin{array}{c:c} \tilde{F} & 0 \\ \hdashline \tilde{C}(t)(\tilde{F}_E - \tilde{F}) & \tilde{A}(t) \end{array}\right] Z(t) + \\ &+ \left[\begin{array}{c:c} \tilde{G} & 0 \\ \hdashline -\tilde{G} & \tilde{C}(t)\tilde{\Gamma}(t)\tilde{L}\tilde{H} \end{array}\right] N(t),\end{aligned} \quad (8)$$

where $Z(t) = \mathrm{vec}\{X(t), \varepsilon(t|t-1)\}$ and $N(t) = \mathrm{vec}\{E(t), V(t)\}$. Applying the mathematical expectation on both sides of (8), we obtain for $\bar{Z}(t) = E\{Z(t)\}$ the recursion

$$\bar{Z}(t+1) = \sum_{i=1}^{\nu} \pi_i B_{[r]} \bar{Z}(t), \quad (9)$$

where $\tilde{B}_{[r]} = \left[\begin{array}{c:c} \tilde{F} & 0 \\ \hdashline \tilde{C}_{[r]}(\tilde{F}_E - \tilde{F}) & \tilde{A}_{[r]} \end{array}\right]$ and $\tilde{C}_{[r]}$ is obtained from $\tilde{C}(t)$ by choosing $\Xi_t = \Xi^{[r]}$.

Similarly, we obtain the following recursion for the mean-square error matrix $P(t) = E\{Z(t)Z(t)^T\}$:

$$P(t+1) = \sum_{i=1}^{\nu} \pi_i [\tilde{B}_{[r]} P(t) \tilde{B}_{[r]}^T + \tilde{D}_{[r]} W \tilde{D}_{[r]}^T], \quad (10)$$

where $\tilde{D}_{[r]} = \left[\begin{array}{c:c} \tilde{G} & 0 \\ \hdashline -\tilde{G} & \tilde{C}_{[r]}\tilde{\Gamma}_{[r]}]\tilde{L}\tilde{H} \end{array}\right]$ and $W = E\{N(t)N(t)^T\} = \mathrm{diag}\{Q^*, \tilde{R}\}$, where $Q^* = \begin{bmatrix} Q & \cdots & Q \\ \vdots & & \\ Q & \cdots & Q \end{bmatrix}$, $\tilde{R} = \mathrm{diag}\{R^{(1)}(t), \ldots, R^{(N)}(t)\}$. Relation (10) can be rewritten in the following way:

$$\begin{aligned}\mathrm{col}\{P(t+1)\} &= \textstyle\sum_{r=1}^{\nu} \pi_r [(\tilde{B}_{[r]} \otimes \tilde{B}_{[r]})\mathrm{col}\{P(t)\} + \\ &+ (\tilde{D}_{[r]} \otimes \tilde{D}_{[r]})\mathrm{col}\{W\}]\end{aligned} \quad (11)$$

where $\mathrm{col}\{.\}$ denotes a vector obtained by concatenating columns of an indicated matrix and the sign $\otimes$ denotes the Kronecker's product.

IV. Stability

For the stability analysis of the proposed estimator, we shall use the following results from the general matrix theory.

Lemma 1. Let $f(.)$ be a matrix norm having the property $f(A) \leq f(B)$ for two $n \times n$ matrices A and B satisfying $A \leq B$ ($A \geq 0$ means that all the elements of A are nonnegative). Let $g(.)$ be any matrix norm and let A be partitioned into square blocks A_{ii}. Then, $h(A)$ is a matrix norm, where

$$h(A) = f\left(\begin{bmatrix} g(A_{11}) & \cdots & g(A_{1k}) \\ \vdots & & \vdots \\ g(A_{k1}) & \cdots & g(A_{kk}) \end{bmatrix}\right). \quad (12)$$

■

Lemma 2. Let A be an $n \times n$ matrix and $\varepsilon > 0$. Then, there exists a matrix norm $\|A\|$ such that

$$\rho(A) \leq \|A\| \leq \rho(A) + \varepsilon, \quad (13)$$

where $\rho(A)$ is the spectral radius of a matrix A ($\rho(A) = \max_i |\lambda_i(A)|$, where $\lambda_i(A)$ are the eigenvalues of A). ■

A norm satisfying the requirement of Lemma 2 is the norm $\|A\|_t = \|D_t U^T A U D_t^{-1}\|_\infty$, where U is an orthogonal matrix in $A = U\Delta U^T$, where Δ is an upper triagonal matrix (Schur's theorem), $D_t = \mathrm{diag}\{t, t^2, t^3, \ldots, t^n\}$ and $\|A\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$ (for $A = [a_{ij}]$, $i, j = 1, \ldots, n$). Inequality (13) is satisfied for any given $\varepsilon > 0$ by choosing $t \geq 0$ large enough.

Lemma 1 can be found in [25] as Conlisk observation, while Lemma 2 and the related statement can be found in [26] (Lemma 5.6.10).

The following theorem gives sufficient conditions for stability in the sense of boundedness of the mean-square error (10). The applied methodology is based on [24], [27] and the definition of a specially constructed norm adapted to the partition of the consensus matrix.

Theorem 1. Let $\tilde{A}_{[r]}$ be partitioned into blocks $\tilde{A}_{jk}^{[r]} = C_{jk}^{[r]} \Phi_j^{[r]}$, where $C_{jk}^{[r]}$ and $\Phi_j^{[r]}$ are realizations of $C_{jk}(t)$ and $\Phi_j(t)$ obtained by choosing $Xi_t = \Xi^{[r]}$, and let $\rho(\Phi_k^{[r]}) < R_k^{[r]}$, $k = 1, \ldots, N$, together with

$$\sum_{r=1}^{\nu} \pi_r (\max_j \sum_{k=1}^{N} \rho(C_{jk}^{[r]}) R_k^{[r]})^2 < 1. \quad (14)$$

Then, the proposed estimator is stable in the sense that $\|E\{\varepsilon(t|t-1)\varepsilon(t|t-1)^T\}\| < \infty$, if the system (1) is stable, or if the system (1) is unstable and $\tilde{F}_E = \tilde{F}$.

Proof: If $\tilde{A}_{[r]}$ is partitioned into $n \times n$ blocks $A_{ij}^{[r]}$, $i, j = 1, \dots, N$, then the matrix $\tilde{A}_{[r]} \otimes \tilde{A}_{[r]}$ is cogredient to $\tilde{A}_{[r]}^P \otimes \tilde{A}_{[r]}^P$, *i.e.* $\tilde{A}_{[r]}^P \otimes \tilde{A}_{[r]}^P = T(\tilde{A}_{[r]} \otimes \tilde{A}_{[r]})T^T$, where T is a permutation transformation, where

$$\tilde{A}_{[r]}^P \otimes \tilde{A}_{[r]}^P =$$

$$\begin{bmatrix} A_{11}^{[r]} \otimes A_{11}^{[r]} & \cdots & A_{11}^{[r]} \otimes A_{1N}^{[r]} & \cdots & A_{1N}^{[r]} \otimes A_{1N}^{[r]} \\ \vdots & & & & \\ A_{21}^{[r]} \otimes A_{11}^{[r]} & \cdots & A_{21}^{[r]} \otimes A_{1N}^{[r]} & \cdots & A_{2N}^{[r]} \otimes A_{1N}^{[r]} \\ \vdots & & & & \\ A_{N1}^{[r]} \otimes A_{N1}^{[r]} & \cdots & A_{N1}^{[r]} \otimes A_{nN}^{[r]} & \cdots & A_{NN}^{[r]} \otimes A_{NN}^{[r]} \end{bmatrix}.$$

Therefore, any norm of $\tilde{A}_{[r]}^P \otimes \tilde{A}_{[r]}^P$ is a norm of $\tilde{A}_{[r]} \otimes \tilde{A}_{[r]}$. Define the norm of $\tilde{A}_{[r]}^P \otimes \tilde{A}_{[r]}^P$ in the following way:

$$\|\tilde{A}_{[r]}^P \otimes \tilde{A}_{[r]}^P\|_c = \left\| \begin{bmatrix} \|A_{11}^{[r]} \otimes A_{11}^{[r]}\|_t & \cdots & \|A_{1N}^{[r]} \otimes A_{1N}^{[r]}\|_t \\ \vdots & & \vdots \\ \|A_{N1}^{[r]} \otimes A_{1N}^{[r]}\|_t & \cdots & \|A_{NN}^{[r]} \otimes A_{NN}^{[r]}\|_t \end{bmatrix} \right\|_\infty$$

where $\|.\|_t$ is defined within Lemma 2. For particular terms in the last expression we have

$$\|A_{jk}^{[r]} \otimes A_{lm}^{[r]}\|_t \le \rho(A_{jk}^{[r]})\rho(A_{lm}^{[r]}) + \varepsilon \le \|A_{jk}^{[r]}\|\|A_{lm}^{[r]}\| + \varepsilon \tag{15}$$

for any $\varepsilon > 0$ and t large enough, according to the Lemma 2. Furthermore, we have $\|A_{jk}^{[r]}\| = \|C_{jk}^{[r]}\Phi_k^{[r]}\|_t \le \rho(C_{jk}^{[r]})\|\Phi_k^{[r]}\|_t$, having in mind that $\|C_{jk}^{[r]}\|_t = \rho(C_{jk}^{[r]})$ for $C_{jk}^{[r]}$ diagonal. Moreover, it is always possible to find such a $\bar{t} > 0$ that for any $t > \bar{t}$ we have $\|\Phi_k^{[r]}\|_t \le \rho(\Phi_k^{[r]}) + \varepsilon$, for any given $\varepsilon > 0$. Making ε small enough we always have $\rho(\Phi_k^{[r]}) + \varepsilon \le R_k^{[r]}$ (having in mind the strict inequality in $\rho(\Phi_k^{[r]}) < R_k^{[r]}$), and, therefore, $\|\Phi_k^{[r]}\|_t \le R_k^{[r]}$. Consequently, one obtains that

$$\|\tilde{A}_{[r]}^P \otimes \tilde{A}_{[r]}^P\|_c \le \max_{j,l}(\sum_{k=1}^{N} \rho(C_{jk}^{[r]})R_k^{[r]})(\sum_{m=1}^{N} \rho(C_{lm}^{[r]})R_m^{[r]}).$$

Coming back to (11), we directly obtain (14). The second conclusion follows trivially from the definition of the matrix $\tilde{B}_{[r]}$. Thus the result. ■

Remark 1. A comparison with the results obtained in relation with the continuous-time estimator based on consensus described in [20] shows basic similarity of the main ideas, but also some differences which should be emphasized. Namely, the main point of the stability analysis in [20] has been to show the existence of stabilizing consensus gains, while the stability conditions given here give a deeper insight into the influence of particular terms, even in the case of intermittent observations and communication faults. ■

Remark 2. The derived conditions are sufficient for stability in the sense of mean-square error boundedness. Following the same line of thought, it is possible to formulate conditions for the convergence of the mean value $\bar{Z}(t)$ to zero (see [19]). It is possible to show that these conditions are based on the modified condition (14): $\sum_{r=1}^{\nu} \pi_r(\max_j \sum_{k=1}^{N} \rho(C_{jk}^{[r]})R_k^{[r]}) < 1$. It can be observed immediately that the last condition is implied by (14), which is logical, having in mind that the requirement for the mean-square error boundedness is stronger that the requirement for the convergence of the mean. ■

Example 1. Assume that the system $\mathbf{S}$ is of first order and unstable, with $F = 1.2$ and that we have two first order local estimators. Assume also that $F_1 = 1.2$, $L_1 = 0.5$ and $H_1 = 1$ for the first estimator, and $F_2 = 1.2$, $L_2 = 0.3$ and $H_2 = 1$ for the second; both estimators are stable when the measurements are available (when $\gamma_i = 1$). The local steady state estimators alone are mean-square stable (in the sense of Theorem 1) if the probabilities p_{11} and p_{22} for not getting measurements are less than $\bar{p}_{11} = 0.6303$ and $\bar{p}_{22} = 0.6741$. Assume now that the network is implemented with constant communication gains $K_{ij} = 0.5$, $i, j = 1, 2$, according to the proposed algorithm (recall (6) and the notation therein). According to Theorem 1, it is possible to find regions in the plane $p_{11} - p_{22}$ which guarantee the mean-square stability of the whole estimator for different values of the communication probability $p = p_{12} = p_{21}$ (Fig. 1, curves with the label (1)). The obtained boundaries are conservative, as expected; however, the beneficial effects of the consensus scheme are obvious. Intercommunications between the agents increase robustness of the local estimators to intermittent measurements. Curves with the label (2) represent, on the same figure, the corresponding boundaries guaranteeing convergence of the mean; their relation with the curves corresponding to the mean-square stability is in accordance with Remark 2. The bounding probabilities for the case of independent estimators are in this case $\bar{p}_{11} = 0.7143$ and $\bar{p}_{22} = 0.7778$, according to (9). ■

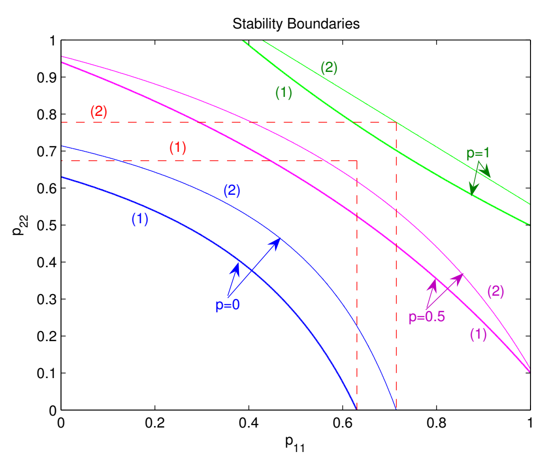

Fig. 1. Stability boundaries

V. Denoising

In this section we shall pay our attention to an important aspect of the analyzed scheme related to its capability to reduce the noise influence by increasing the number of active measuring and processing agents. The basic problem of consensus averaging has been studied for different network topologies in, *e.g.*, [17]. In [20] and [21], an analysis is given of the continuous-time Kalman filtering algorithms and stochastic approximation algorithms connected by a dynamic consensus scheme, which demonstrates that it is possible to reduce the measurement noise influence by increasing the complexity of the network. We shall provide here an insight into this problem in the context of the proposed discrete-time state estimation algorithm by using a special simplified setting which enables getting qualitatively clear conclusions. Generalizations to more complex structures are feasible, although technically more complex.

We shall assume further that all the estimators have the information about the system model, and that they observe the identical components of the state vector, but with different realizations of the measurement noise, having the same covariance R. We shall also assume that the measurements are never interrupted, and that there are no communication faults.

Case (a): the consensus matrix $\tilde{C}(t)$ is constant and given in the form $\tilde{C}(t) = \tilde{C}_1^{(N)} = \frac{1}{N}\begin{bmatrix} I & I & \cdots & I \\ I & I & \cdots & I \\ & & \cdots & \\ I & I & \cdots & I \end{bmatrix}$, where I stands for I_n.

From (10) we obtain for the steady-state estimation mean-square error S the following Lyapunov-like algebraic equation:

$$S = \sum_{r=1}^{\nu} \pi_r [\tilde{A}_{[r]}^N S \tilde{A}_{[r]}^T + \tilde{E}_{[r]} W \tilde{E}_{[r]}^T], \tag{16}$$

where $\tilde{E}_{[r]} = \left[-\tilde{G} \,\vdots\, \tilde{C}_{[r]} \tilde{\Gamma}_{[r]}] \tilde{L} \tilde{H} \right]$. The above simplifying assumptions lead to

$$S^{(N)} = \tilde{C}_1^{(N)} [\tilde{\Phi} S^{(N)} \tilde{\Phi}^T + \tilde{L} \tilde{H} \tilde{R} \tilde{H}^T \tilde{L}^T] \tilde{C}_1^{(N)} + \tilde{G} Q^* \tilde{G}^T \tag{17}$$

where the upper index $^{(N)}$ emphasizes that there are N agents; the block-diagonal matrices $\tilde{\Phi}$, $\tilde{L}$, $\tilde{H}$ and $\tilde{G}$ are composed of identical block-diagonal elements.

We observe now that $\tilde{C}_1^{(N)}$ has n eigenvalues at 1, and $(N-1)n$ eigenvalues at 0. Its diagonalization can be done using

$$T_N = \begin{bmatrix} I & I & \cdots & I \\ I & -(N-1)I & \cdots & I \\ & & \cdots & \\ I & & \cdots & -(N-1)I \end{bmatrix},$$

so that

$$T_N^{-1} \tilde{C}_1^{(N)} T_N = \bar{C}_1^{(N)} = \left[\begin{array}{c:ccc} I & 0 & \cdots & \\ \hdashline 0 & 0 & 0 & \cdots \\ & & \cdots & \\ 0 & 0 & \cdots & 0 \end{array}\right]. \tag{18}$$

Applying T_N^{-1} and T_N to equation (17), we obtain

$$\bar{S}^{(N)} = \bar{C}_1^{(N)} [\tilde{\Phi} \bar{S}^{(N)} \tilde{\Phi}^T + \tilde{L} \tilde{H} \tilde{R} \tilde{H}^T \tilde{L}^T] \bar{C}_1^{(N)} + \bar{Q}^{(N)} \tag{19}$$

where $\bar{S}^{(N)} = T_N^{-1} S^{(N)} T_N$ and $\bar{Q}^{(N)} = T_N^{-1} \tilde{G} Q^* \tilde{G}^T T_N = \begin{bmatrix} NGQG^T & \cdots & 0 \\ \vdots & & \\ 0 & \cdots & 0 \end{bmatrix}$. A solution to this equation is $\bar{S}^{(N)} = \left[\begin{array}{c:ccc} \hat{S}^{(N)} & 0 & \cdots & \\ \hdashline 0 & 0 & 0 & \cdots \\ & & \cdots & \\ 0 & 0 & \cdots & 0 \end{array}\right]$, where $\hat{S}^{(N)}$ is obtained from the Lyapunov equation

$$\hat{S}^{(N)} = \Phi \hat{S}^{(N)} \Phi^T + LHRH^T L^T + NGQG^T. \tag{20}$$

Obviously, the mean-square error for the whole vector $X(t)$ is $J = \operatorname{Tr} \bar{S}^{(N)} = \operatorname{Tr} \hat{S}^{(N)}$. Having in mind that N independent estimators have the mean-square error equal to $N\hat{J}$, where $\hat{J} = \operatorname{Tr} \hat{S}$ and $\hat{S}$ is a solution to the standard Lyapunov equation

$$\hat{S} = \Phi \hat{S} \Phi^T + LHRH^T L^T + GQG^T, \tag{21}$$

we take $\bar{J} = \frac{1}{N} J$ as the average criterion "per agent", and obtain that for N large enough $\bar{J} \approx \operatorname{Tr} S^*$, where S^* is a solution of the Lyapunov equation

$$S^* = \Phi S^* \Phi^T + GQG^T. \tag{22}$$

Comparing (21) and (22), one concludes that for large N the consensus scheme asymptotically reduces the mean square error from the level defined by (21) to the level defined by (22), where the measurement noise is reduced to zero. In this sense, introduction of the inter-agent communications based on consensus contributes to the measurement noise suppression, that is, to *denoising*.

Our previous case was related to the network with the maximal connectivity: the next case is related to the network with minimal connectivity.

Case (b): $\tilde{C}(t) = \tilde{C}_2^{(N)} = \frac{1}{2}\begin{bmatrix} I & I & 0 & \cdots & 0 \\ 0 & I & I & 0 & \cdots \\ & & \cdots & & \\ 0 & 0 & \cdots & I & I \\ I & 0 & \cdots & 0 & I \end{bmatrix}$, *i.e.*, the network graph forms a directed ring.

Diagonalization of $\tilde{C}_2^{(N)}$ produces a diagonal matrix with the eigenvalues $\lambda_i^{(N)}$, $i = 1, \ldots, N$, uniformly distributed on a circle in the complex plane of radius $\frac{1}{2}$ with the center at $(\frac{1}{2}, 0)$. Consequently, we get for the diagonal blocks $(\bar{S}^{(N)})_i$, $i = 2, \ldots, N$, the following relations:

$$(\bar{S}^{(N)})_i = |\lambda_i^{(N)}|^2 [\Phi (\bar{S}^{(N)})_i \Phi^T + LHRL^T H^T], \tag{23}$$

except for the first one, which is the same as in Case (a). It is obvious that in this case the sum of the terms in (23) does not tend to zero, so that denoising in the above sense is not achievable in spite of the fact that all the nodes are reachable from any other node.

However, the relation (23) indicates how asymptotic denoising can be achieved in the case of graphs with complexity lying between the above two extremes. Summing up the terms defined by (23), we easily conclude that

$$\|\Sigma_{i=2}^N (\bar{S}^{(N)})_i\| \leq k \sum_{i=1}^n |\lambda_i^{(N)}|^2$$

for some finite $k > 0$. Therefore, the condition

$$\sum_{i=1}^{N} |\lambda_i^{(n)}|^2 = o(N) \tag{24}$$

is sufficient for successful asymptotical denoising.

Example 2. This example illustrates the denoising capabilities of the proposed estimator. We assume that all the agents have the identical models of the stable fourth order system with $Q = 0.5I_4$, $R = 0.5I_4$, but with different realizations of the measurement noise. It is also assumed that the measurements are never interrupted and that there are no communication faults (with probability 1). Average values of the criterion ($\bar{J}$) have been calculated for two network topologies: a) fully connected network; b) directed ring. The results are shown in Fig 1. The horizontal dashed line correspond to the criterion lower bound $\mathrm{Tr}(S^*)$ where S^* is obtained by using (22). The presented results fully confirm the above analysis. In the case of the fully connected graph, as the number of agents (N) goes to infinity the curve exactly converges to the lower bound of the criterion. In the case of the directed ring the limit value of the criterion is higher than $\mathrm{Tr}(S^*)$; hence, the complete asymptotic denoising is not achieved.

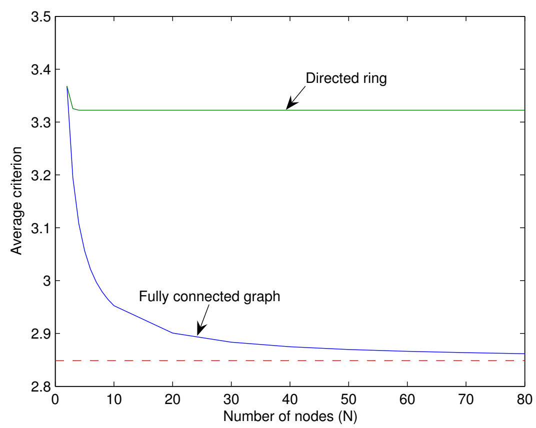

Fig. 2. Average criterion as a function of N

VI. Conclusion

In this paper a new algorithm for overlapping decentralized state estimation has been considered. It is based on a synergy between the local Kalman filters, with possible intermittent observations, and a consensus scheme which forms a communication network between the agents, with possible communication errors. Using a specially defined matrix norm, sufficient conditions for the asymptotic stability in the sense of bounded mean-square estimation error have been derived. The important problem of denoising has also been considered. It has been shown that the algorithm is able to contribute to the measurement noise suppression by increasing the network complexity.

References

[1] C. W. Sanders, E. C. Tacker, and T. D. Linton, "A new class of decentralized filters for interconnected systems," *IEEE Trans. Autom. Control*, vol. 19, pp. 259–262, 1974.

[2] ——, "Specific structures for large scale state estimation algorithms having information exchange," *IEEE Trans. Autom. Control*, vol. 23, pp. 255–260, 1978.

[3] D. D. Šiljak, *Decentralized Control of Complex Systems*. New York: Academic Press, 1991.

[4] J. L. Speyer, "Computation and transmission requirements for a decentralized linear quadratic gaussian control problem," *IEEE Trans. Autom. Control*, vol. 24, pp. 266–269, 2004.

[5] E. C. Tacker and C. W. Sanders, "Decentralized structures for state estimation in large scale systems," *Large Scale Systems*, vol. 40, pp. 39–49, 1980.

[6] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Prentice Hall, 1989.

[7] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Dep. Electrical Eng. Comput. Sci., M.I.T., Cambridge, MA, 1984.

[8] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. Autom. Control*, vol. 31, pp. 803–812, 1986.

[9] A. Fax and R. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Trans. Automat. Contr.*, vol. 49, pp. 1465–1476, 2004.

[10] A. Jadbabaie, J. Lin, and A. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Automat. Contr.*, vol. 48, pp. 988–1001, 2003.

[11] Z. Lin, B. Francis, and M. Maggiore, "Necessary and sufficient conditions for formation control of unicycles," *IEEE Trans. Automat. Contr.*, vol. 50, pp. 121–127, 2005.

[12] L. Moreau, "Stability of multiagent systems with time-dependent communication links," *IEEE Trans. Automat. Contr.*, vol. 50, pp. 169–182, 2005.

[13] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Automat. Contr.*, vol. 49, pp. 1520–1533, 2004.

[14] W. Ren and R. Beard, "Consensus seeking in multi-agent systems using dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, pp. 655–661, 2005.

[15] W. Ren, R. W. Beard, and D. B. Kingston, "Multi-agent Kalman consensus with relative uncertainty," in *Proc. American Control Conference*, 2005.

[16] R. Olfati-Saber, "Distributed Kalman filter with embedded consensus filters," in *Proc. IEEE Conf. Decision and Control*, 2005.

[17] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems and Control Letters*, vol. 53, pp. 65–78, 2004.

[18] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. Int. Conf. Inf. Proc. in Sensor Networks*, 2005.

[19] S. S. Stanković, M. S. Stanković, and D. M. Stipanović, "Consensus based overlapping decentralized estimation with missing observations and communication faults," - to be presented at the IFAC Congress, 2008.

[20] ——, "Consensus based overlapping decentralized estimator," in *Proc. American Control Conference*, 2007.

[21] ——, "Decentralized parameter estimation by consensus based stochastic approximation," in *Proc. IEEE Conference on Decision and Control*, 2007.

[22] S. S. Stanković and D. D. Šiljak, "Stabilization of fixed modes in expansions of LTI systems," - to be published in *Systems and Control Letters*, 2007.

[23] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. New York: Prentice Hall, 1979.

[24] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poola, M. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, pp. 1453–1464, 2004.

[25] I. F. Pierce, "Matrices with dominating diagonal blocks," *Journ. of Economic Theory*, vol. 9, pp. 159–170, 1974.

[26] R. A. Horn and C. A. Johnson, *Matrix Analysis*. Cambridge Univ. Press, 1985.

[27] J. Nilsson, "Analysis and design of real-time systems with random delays," Ph.D. dissertation, Lund Institute of Technology, 1996.