State estimation in a centralized sensor network under limited communication rate

Giorgio Battistelli
Dip.to di Sistemi e Informatica
Università di Firenze
50139 Firenze, Italy
Email: battistelli@dsi.unifi.it

Alessio Benavoli
Istituto "Dalle Molle"
di Studi sull'Intelligenza Artificiale,
6928 Manno-Lugano, Switzerland,
Email: alessio@idsia.ch

Luigi Chisci
Dip.to di Sistemi e Informatica
Università di Firenze
50139 Firenze, Italy
Email: chisci@dsi.unifi.it

***Abstract*— This paper deals with the problem of estimating the state of a discrete-time linear stochastic dynamical system on the basis of data collected from multiple sensors subject to a limitation on the communication rate from the remote sensor units. More specifically, the attention is devoted to a centralized sensor network consisting of: (1) S remote nodes which collect measurements of the given system, compute state estimates at the full measurement rate and transmit them at a reduced communication rate; (2) a fusion node F that, based on received estimates, provides an estimate of the system state at the full rate. A measurement-independent strategy for deciding when transmitting estimates from each sensor to F will be considered. Sufficient conditions for the boundedness of the state covariance at node F will be given. Further, the possibility of determining a communication strategy with optimal performance in terms of minimum mean square estimation error will be investigated.**

I. Introduction

This paper deals with the problem of estimating the state of a discrete-time linear stochastic dynamical system

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{w}_k \tag{I.1}$$

on the basis of measurements collected from multiple sensors

$$\mathbf{y}_k^i = \mathbf{C}^i\mathbf{x}_k + \mathbf{v}_k^i \tag{I.2}$$

for $i = 1, \ldots, S$ subject to a limitation on the communication rate from each remote sensor unit to the state estimation unit. More specifically, the attention will be focused on the use of a centralized sensor network (see Fig. 1) consisting of:

- S remote *sensing nodes* $1, \ldots, S$ which collect noisy measurements $\mathbf{y}_k^1, \ldots, \mathbf{y}_k^S$ of the given system, can process them to find filtered estimates $\hat{\mathbf{x}}_{k|k}^1, \ldots, \hat{\mathbf{x}}_{k|k}^S$ and transmit such estimates to the fusion node;
- a local *fusion node* F which receives data from the S sensors and, based on such data, should provide, in the best possible way, an estimate $\hat{\mathbf{x}}_{k|k}$ of the system's state.

It is assumed that for each sensor (i, $i = 1, \ldots, S$) the ratio

$$\frac{\text{communication rate } i \rightarrow F}{\text{measurement rate of } i} \; = \; \alpha^i \in (0, 1)$$

where possibly $\alpha^i \ll 1$. Hereafter, without loss of generality, the unit rate will be fixed equal to the measurement rate so that α^i actually denotes the communication rate.

In particular, the objective is to devise for each sensor a *communication strategy* with fixed rate α^i that guarantees bounded state covariance and possibly optimal estimation performance, in terms of minimum *Mean Square Error (MSE)* at the node F.

The above scenario reflects the practical situation in which the sensing units and the monitoring unit are remotely dislocated with respect to each other and the communication rate between them is severely limited by energy, band and/or security concerns.

State estimation under finite communication bandwidth has been thoroughly investigated, see e.g. [1]-[6]. In the above cited references, the emphasis is on the analysis of the quantization effects due to the encoding of transmitted data into a finite alphabet of symbols as well as on the design of efficient, possibly optimal, coding algorithms. Conversely, following [7], [8] the present work tackles the issue of communication bandwidth finiteness from a completely different viewpoint. Specifically, it is assumed that infinite-precision data are transmitted over the communication channels[1] while the bandwidth limitation is accomplished by imposing suitable values of the communication rates α^i. In this context, the focus will be on the choice of a communication strategy for deciding which data transmit from each sensor i to F. In particular, measurement-independent strategies will be considered and issues related to the stability and optimality of such strategies will be investigated. Two main results will be provided extending to centralized sensor networks the results of [8] obtained in the case of a single remote sensing unit. First it will be shown that boundedness of the MSE can be ensured provided that an upper bound on the time between consecutive transmissions is imposed. Secondly, an optimization-based approach will be proposed to derive the communication strategy with minimum MSE. The proofs are omitted due to space limitations

The notations are quite standard: $\mathbb{Z}_+$ is the set of nonnegative integers; given a square matrix $\mathbf{M}$, $\mathrm{tr}(\mathbf{M})$ denotes its trace; $\mathbb{E}\{\cdot\}$ and $\mathbb{P}\{\cdot\}$ denote the expectation and, respectively, probability operators; finally, given a generic sequence $\{s_k;\, k = 0, 1, \ldots\}$ and two time instants $k_1 \leq k_2$, we define $s_{k_1:k_2} \triangleq \{s_{k_1}, s_{k_1+1}, \ldots, s_{k_2}\}$.

[1]This assumption, though incompatible with the finite bandwidth, holds in practice provided that quantization errors are negligible with respect to measurement errors.

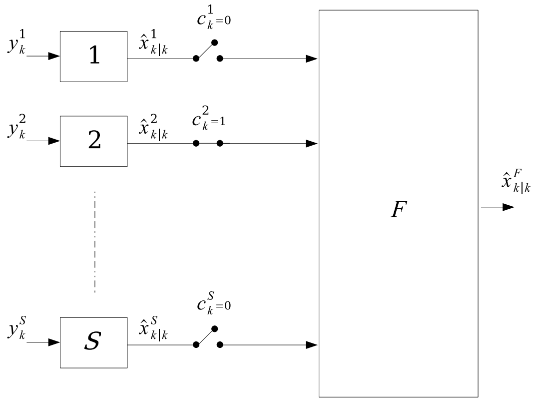

Fig. 1. The considered centralized sensor network.

II. Communication strategy

The aim of this section is to formalize the concept of *communication strategy (CS)* with fixed rate α^i for sensor i. To this end, let us introduce for each sensor i binary variables c_k^i such that

$$c_k^i \;=\; \begin{cases} 1, & \text{if sensor } i \text{ transmits at time } k \\ 0, & \text{if sensor } i \text{ does not transmit at time } k \end{cases}$$

The communication strategy is characterized by the following two choices:

- the type of data being transmitted by each sensor i, either measurements $\mathbf{y}_k^i$ or filtered state estimates $\hat{\mathbf{x}}^i_{k|k}$;
- the mechanism by which sensor i decides whether to transmit or not at time k, i.e. the mechanism of generating $c_k^i \in \{0,1\}$.

As far as the first choice is concerned, the following options can for instance be adopted:

- **measurement transmission**: when $c_k^i = 1$, sensor i transmits $\mathbf{y}_k^i$ to the fusion node F;
- **estimate transmission**: when $c_k^i = 1$, sensor i transmits $\hat{\mathbf{x}}^i_{k|k}$ to the fusion node F.

In this paper, the attention will be restricted to estimate transmission, assuming that each sensor i has enough processing capability to update on-line some state estimate $\hat{\mathbf{x}}^i_{k|k}$. Note that, with such a choice, there is no loss of information due to the finite bandwidth as the estimate $\hat{\mathbf{x}}^i_{k|k}$ represents a sufficient statistics.

As far as the decision mechanism is concerned, this can be formally defined as follows.

Definition - A decision mechanism with rate $\alpha^i \in (0,1)$ is any, deterministic or stochastic, mechanism of generating c_k^i such that

$$\lim_{t\to\infty} \frac{1}{t} \sum_{k=1}^{t} \mathbb{E}\left\{c_k^i\right\} \;=\; \alpha^i \qquad \text{(II.1)}$$

Several decision mechanisms can clearly be devised. In this work, the attention will be restricted to *measurement-independent* strategies that, at time k, decide whether to transmit or not independently of the measurement sequence $\mathbf{y}^i_{0:k}$. More specifically, it is supposed that each c_k^i is chosen according to some probabilistic criterion, possibly adapted only on the basis of the past transmission pattern, so as to ensure the desired communication rate. With this respect, let $n_k^i \geq 0$ denote, at a generic time k, the number of time instants elapsed from the last transmission of sensor i, i.e.

$$c^i_{k-n_k^i} = 1 \;\text{ and }\; c_k^i = c^i_{k-1} = \cdots = c^i_{k-n_k^i+1} = 0.$$

Note that n_k^i is a function of $c^i_{0:k}$ and can be recursively computed as

$$n_k^i = (1 - c_k^i)\,(n^i_{k-1} + 1)\,.$$

It should be evident that, in the considered framework, such a quantity provides a rough estimate of how good is the estimate $\hat{\mathbf{x}}^i_{k|k-n_k^i}$ available at the node F. Then it seems reasonable to take into account the value of n^i_{k-1} when choosing whether to transmit or not at the generic time k. In this connection, the conditional probabilities

$$\mathbb{P}\{c_k^i = 1 | n^i_{k-1} = j\}\,, \quad j = 0, 1, \ldots$$

can be considered as design parameters in the communication strategy that can be suitably tuned to improve the performance (e.g., to reduce the MSE). This will be the subject of Sections III and IV.

Throughout the paper, the following notation will be adopted

$$\begin{aligned} \mathbf{c}_k &\triangleq \operatorname{col}\left(c_k^1, \ldots, c_k^S\right), \quad \mathbf{n}_k \triangleq \operatorname{col}\left(n_k^1, \ldots, n_k^S\right), \\ \mathbf{y}_k &\triangleq \operatorname{col}\left(\mathbf{y}_k^1, \ldots, \mathbf{y}_k^S\right). \end{aligned}$$

III. Transmission strategy and boundedness of the error covariance

In this section, the focus is on transmission strategies wherein, for each sensor $i \in \{1, \ldots, S\}$, the binary variable c_k^i is a probabilistic function of the number n_k^i of time instants elapsed from the last transmission of sensor i.

Specifically, at the generic time k, c_k^i is chosen to be a Bernoulli random variable with parameter $\varphi^i(n^i_{k-1})$, i.e., c_k^i takes value 1 with probability $\varphi^i(n^i_{k-1})$ and value 0 with probability $1 - \varphi^i(n^i_{k-1})$. Clearly, this corresponds to the assignments

$$\mathbb{P}\{c_k^i = 1 | n^i_{k-1} = j\} = \varphi^i(j)\,,$$

for $j = 0, 1, \ldots$ and $i = 1, \ldots, S$. The functions $\varphi^i : \mathbb{Z}_+ \to [0,1]$ must be chosen so that the transmission rate constraint is met for each sensor i.

Remark 1: It is important to point out that the dependence of the transmission strategy on the elapsed times $\mathbf{n}_k$ can be crucial to ensure that the estimation error is eventually bounded in mean square. With this respect, in [8] it has been shown that, in the case of a single remote sensor, when $\varphi(j) = \alpha$ for $j = 0, 1, \ldots$ the MSE at node F may diverge if the transmission rate α does not exceed a certain critical value.

Nevertheless, it turns out that boundedness of the MSE can be ensured provided that an upper bound on the time between consecutive transmissions is imposed. Then, attention will be devoted to transmission strategies satisfying the following assumption.

A1. The transmission strategy is such that for each sensor $i \in \{1, \ldots, S\}$, one has $\varphi^i(N^i - 1) = 1$ for some $N^i > 1$.

This amounts to assuming that the time interval between two consecutive transmissions of sensor i never exceeds N^i. Clearly, assumption A1 implies that the vector of elapsed times $\mathbf{n}_k$ always takes value in the set

$$\mathcal{N} \stackrel{\triangle}{=} \left\{ (n^1, \ldots, n^S)' \in \mathbb{Z}_+^S : n^i \le N^i - 1\,,\ i = 1, \ldots, S \right\}.$$

As it will be shown in the next section, transmission strategies of this kind satisfying the communication rate constraint (II.1) always exist provided that $N^i > 1/\alpha^i$. It is also worth noting that such a class of strategies comprises periodic strategies with transmission rate equal to $1/N^i$ (in this case $\varphi^i(j) = 0$, for $j = 0, \ldots, N^i - 2$).

As well known, different fusion algorithms can be devised to combine data collected from multiple sensors and related information in order to derive a fused estimate $\hat{\mathbf{x}}_{k|k}$ of the true state $\mathbf{x}_k$. As to the computation of the MSE corresponding to a generic transmission strategy, some hypotheses about the fusion algorithm are needed. Specifically, it is supposed that the fusion algorithm has the following properties.

A2. At each time $k = 0, 1, \ldots$, the fused estimate $\hat{\mathbf{x}}_{k|k}$ is unbiased.

A3. At each time $k = 0, 1, \ldots$, the covariance matrix of the estimation error $\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}$ is a function of the last transmission instant of each sensor, i.e.,

$$\mathbb{E}\left\{ (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})' \,|\, \mathbf{y}_{0:k}\,;\, \mathbf{c}_{0:k} \right\} = \mathbf{P}_{k|k}(\mathbf{n}_k)\,.$$

A4. For each $\mathbf{n} \in \mathcal{N}$, the covariance matrix $\mathbf{P}_{k|k}(\mathbf{n})$ converges exponentially to a steady state value $\mathbf{P}(\mathbf{n})$.

Assumptions A2-A4 can be seen as a sort of minimal requirements that a sensible information fusion algorithm should satisfy. In Section V a fusion algorithm based on a BLUE approach will be proposed satisfying all such properties.

By exploiting assumptions A1 and A3, one can write the mean covariance of the estimation error as

$$\mathbb{E}\left\{ (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})' \right\} = \sum_{\mathbf{n} \in \mathcal{N}} \mathbb{P}\{\mathbf{n}_k = \mathbf{n}\} \mathbf{P}_{k|k}(\mathbf{n})\,.$$

Since in the considered framework the sensors' transmission strategies are independent, then for each vector $\mathbf{n} = (n^1, \ldots, n^S)$ the probability $\mathbb{P}\{\mathbf{n}_k = \mathbf{n}\}$ can be decomposed as

$$\mathbb{P}\{\mathbf{n}_k = \mathbf{n}\} = \prod_{i=1}^{S} \mathbb{P}\{n_k^i = n^i\}\,.$$

Moreover, note that each sequence $n_0^i, n_1^i, \ldots$ can be described by a discrete-time Markov chain characterized by the state set $\{0, 1, \ldots, N^i - 1\}$ and the transition matrix

$$\mathbf{\Phi^i} = \begin{bmatrix} \varphi^i(0) & \varphi^i(1) & \ldots & \varphi^i(N^i-2) & 1 \\ 1-\varphi^i(0) & 0 & \ldots & 0 & 0 \\ 0 & 1-\varphi^i(1) & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1-\varphi^i(N^i-2) & 0 \end{bmatrix}$$

For the reader's convenience, the associated state transition diagram is provided in Fig. 2. It is easy to check that the Markov chain $n_0^i, n_1^i, \ldots$ is irreducible if and only if $\varphi^i(j) < 1$ for $j = 0, \ldots, N^i - 2$. In this case, there exists a unique invariant distribution for the probabilities $\mathbb{P}\{n_k^i = j\} = p^i(j)\,,\quad j = 0, \ldots, N^i - 1$, that represents also the long-run average distribution, i.e.,

$$\lim_{t \to \infty} \frac{1}{t} \sum_{k=1}^{t} \mathbb{P}\{n_k^i = j\} = p^i(j)\,. \tag{III.1}$$

As well known, two possibilities may arise:

(a) the Markov chain is aperiodic and each probability $\mathbb{P}\{n_k^i = j\}$ converges exponentially to $p^i(j)$;

(b) the Markov chain is periodic and its asymptotic behavior depends on the initial probabilities (condition (III.1) holds only on average).

In the context of our application, it is desirable that in the long run the probabilities $\mathbb{P}\{n_k^i = j\}$ do not depend on their initial values, otherwise the error covariance at node F would depend on the phase shift among the sensors and a synchronization would be needed to optimize the performance. Then the following constraint is imposed.

A5. For each sensor $i = 1, \ldots, S$, the initial condition n_0^i is chosen according to the steady-state probability distribution $\mathbb{P}\{n_0^i = j\} = p^i(j)\,,\quad j = 0, \ldots, N^i - 1\,.$

Clearly, this implies that each probability $\mathbb{P}\{n_k^i = j\}$ coincides with $p^i(j)$ for any $k = 0, 1, \ldots$.

For what concerns the invariant distribution $\mathbf{p}^i \stackrel{\triangle}{=} [p^i(0) \quad \ldots \quad p^i(N^i - 1)]'$, the following result can be readily established.

Lemma 1: For each sensor i, the invariant distribution is obtained as

$$p^i(j) = \left(\sum_{j=0}^{N^i-1} \prod_{l=0}^{j-1} (1 - \varphi^i(l)) \right)^{-1} \prod_{l=0}^{j-1} (1 - \varphi^i(l))\,, \tag{III.2}$$

for $j = 0, \ldots, N^i - 1\,.$ □

It is important to remark that, when $\varphi^i(j) = 1$ for some $j < N^i - 1$, the considered Markov chain is not irreducible. However, in this case, there exists a unique closed communicating class corresponding to the set of states $\{0, 1, \ldots, M^i - 1\}$, where M^i is the smallest integer such that $\varphi^i(M^i) = 1$. All the other states have null asymptotic

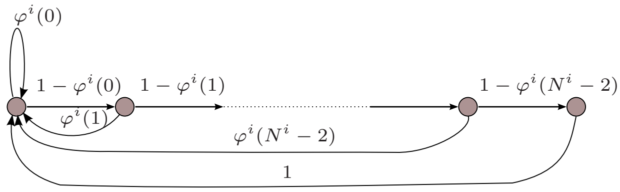

Fig. 2. State transition diagram of the Markov chain $n_0^i, n_1^i, \ldots$ for the modified strategy.

probabilities. Then, it is straightforward to see that (III.1) and (III.2) still hold. This generalization will be useful in the next section.

The foregoing discussion leads to the following result.

Theorem 1: Suppose that assumptions A1-A5 hold. Then the mean covariance of the estimation error can be obtained as

$$\lim_{t\to\infty} \frac{1}{t}\sum_{k=1}^{t} \mathbb{E}\left\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})'\right\} = \sum_{(n^1,\ldots,n^S)\in\mathcal{N}} \left(\prod_{i=1}^{S} p^i(n^i)\right) \mathbf{P}\left((n^1,\ldots,n^S)'\right) .$$

□

Note that, such a covariance is always bounded regardless of the transmission rates α^i and of the stability of the system matrix $\mathbf{A}$.

IV. Optimized transmission strategies

This section is devoted to the synthesis of an optimized transmission strategy. In this connection, note that, for given $N^1,\ldots,N^S$, the strategy introduced in the previous section is not uniquely determined in that there are $\sum_{i=1}^{S}(N^i-1)$ degrees of freedom, i.e., the parameters $\varphi^i(j)$ for $j=0,\ldots,N^i-2$ and $i=1,\ldots,S$. To be more precise, when the transmission rates α^i are fixed, the degrees of freedom are only $\sum_{i=1}^{S}(N^i-2)$ since the parameters have to be chosen so that the equality constraint (II.1) is satisfied for each sensor. It is quite natural to think of exploiting such degrees of freedom in order to optimize some performance index. With this respect, a reasonable choice for the performance index is given by the cost

$$J \triangleq \lim_{t\to\infty} \frac{1}{t}\sum_{k=1}^{t} \mathrm{tr}\left(\mathbb{E}\left\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})'\right\}\, \mathbf{W}\right)$$

where $\mathbf{W}$ is a positive definite weight matrix.

In the light of Theorem 1, it is immediate to write cost J as a function of the parameters $\varphi^i(j)$ for $j=0,\ldots,N^i-2$ and $i=1,\ldots,S$:

$$J = \sum_{(n^1,\ldots,n^S)\in\mathcal{N}} \mathrm{tr}\left(\mathbf{P}((n^1,\ldots,n^S)')\mathbf{W}\right)\left(\prod_{i=1}^{S} p^i(n^i)\right) \tag{IV.1}$$

where the steady state probabilities $p^i(n^i)$ can be obtained as in (III.2). In the sequel, for the sake of compactness, we shall use the notation $w(n^1,\ldots,n^S) \triangleq \mathrm{tr}\left(\mathbf{P}((n^1,\ldots,n^S)')\mathbf{W}\right)$.

As to the transmission rate constraint (II.1), the following result holds.

Lemma 2: For each sensor i, the transmission rate constraint (II.1) is equivalent to

$$\sum_{j=0}^{N^i-1}\prod_{l=0}^{j-1}(1-\varphi^i(l)) \;=\; 1/\alpha^i\,. \tag{IV.2}$$

□

Summing up, the following constrained optimization problem can be stated that leads to the synthesis of an optimized strategy.

Problem 1: Given $N^1,\ldots,N^S$, find the optimal parameters $\varphi^i(j)$ for $j=0,\ldots,N^i-2$ and $i=1,\ldots,S$ that minimize cost J under the constraints

$$\begin{aligned}
&\sum_{j=0}^{N^i-1}\prod_{l=0}^{j-1}(1-\varphi^i(l)) \;=\; 1/\alpha^i\,, && i=1,\ldots,S\\
&0\le\varphi^i(j)\le 1\,, && j=0,\ldots,N^i-2,\\
& && i=1,\ldots,S\,.
\end{aligned}$$

□

In order to address the solution of Problem 1, it is convenient to recast it in terms of the steady state probabilities $p^i(n^i)$. To this end, note that each constraint (IV.2) corresponds to $p^i(0)=\alpha^i$. Further, if one refers to (III.2), the probabilities $p^i(j)$ can also be obtained recursively as

$$p^i(j) = p^i(j-1)(1-\varphi^i(j-1))\,,\quad j=1,\ldots,N^i-1\,.$$

Then the feasibility constraints $0\le\varphi^i(j)\le 1$, $j=0,\ldots,N^i-2$, turn out to be equivalent to

$$\begin{aligned}
&p^i(j)\ge 0\,, && j=1,\ldots,N^i-1\,,\\
&p^i(j+1)\le p^i(j)\,, && j=0,\ldots,N^i-2\,.
\end{aligned} \tag{IV.3}$$

Finally, notice that $\varphi^i(N)=1$ implies that the probabilities $p^i(j)$, $j=0,\ldots,N^i-2$, sum up to one, i.e.,

$$\sum_{j=0}^{N^i-1} p^i(j) \;=\; 1\,. \tag{IV.4}$$

As a consequence, each vector $\mathbf{p}^i$ is constrained to belong to the polytope $\mathcal{P}^i$ defined by (IV.3), (IV.4) and $p^i(0)=\alpha^i$. Thus, Problem 1 admits the equivalent formulation:

$$\begin{aligned}
&\text{minimize} && \sum_{(n^1,\ldots,n^S)\in\mathcal{N}} w(n^1,\ldots,n^S)\left(\prod_{i=1}^{S} p^i(n^i)\right)\\
&\text{subject to} && \mathbf{p}^i\in\mathcal{P}^i\,,\qquad i=1,\ldots,S\,.
\end{aligned} \tag{IV.5}$$

In the light of such a reformulation, the following feasibility result can be readily established.

Proposition 1: Problem 1 admits solution if and only if $N_i \ge 1/\alpha^i$ for $i=1,\ldots,N$. □

For what concerns the solution of (IV.5), it is worth noting that for any i the objective function $J(\mathbf{p}^1,\ldots,\mathbf{p}^S)$ is linear in $\mathbf{p}^i$ when all $\mathbf{p}^j$, $j \neq i$, are fixed. Thus solving (IV.5) amounts to finding the minimum of a multilinear function over the Cartesian product of Polytopes $\mathcal{P} \triangleq \mathcal{P}^1 \times \cdots \times \mathcal{P}^S$. Then, at least in principle, one can exploit the known fact that there always exists an extreme point of $\mathcal{P}$ where the multilinear function is minimized and solve the problem in finite time by means of an extensive search over all the extreme points of $\mathcal{P}$. The main drawback of such an approach is that the number of extreme points of $\mathcal{P}$ grows exponentially with the number of sensors S. As a consequence, unless very small instances of the problem are considered, the possibility of solving it exactly is ruled out by the so-called *curse of dimensionality*, i.e., the exponential growth of the computational burden. However, it is possible to exploit the particular structure of the problem and devise an iterative optimization algorithm that converges to a local minimum. For example one may apply the following variant of the well-known cyclic coordinate descent.

Algorithm 1:

1) choose a feasible starting point $(\mathbf{p}^{1,0}, \ldots, \mathbf{p}^{S,0})$ and set $k = 0$;
2) for $i = 1, \ldots, S$ find a vector $\mathbf{p}^{i,k+1} \in \mathcal{P}^i$ that minimizes
$$J(\mathbf{p}^{1,k+1}, \ldots, \mathbf{p}^{i-1,k+1}, \mathbf{p}^i, \mathbf{p}^{i+1,k}, \ldots, \mathbf{p}^{S,k})$$
with respect to $\mathbf{p}^i$;
3) if $J(\mathbf{p}^{1,k+1}, \ldots, \mathbf{p}^{S,k+1}) = J(\mathbf{p}^{1,k}, \ldots, \mathbf{p}^{S,k})$ then stop, otherwise set $k = k + 1$ and go to step 1.

A few remarks about Algorithm 1 are in order. First note that each minimization in step 2 is a linear program of polynomial size, which can be solved in polynomial time and space using either the ellipsoid algorithm or the interior-point method [9]. Further, in spite of an exponential worst-case complexity, the simplex algorithm can solve linear programs quite efficiently in practice with average complexity $O(\min\{n^2, m^2\})$ for a problem with n variables and m constraints [10]. Further, when the termination criterion in step 3 is satisfied, the algorithm has converged to a local minimum representing a Nash equilibrium in that no sensor i can decrease the objective function by changing unilaterally its probabilities vector $\mathbf{p}^i$.

V. Fusion algorithm

This section is devoted to the description of the operations that are performed in the sensor nodes $1, \ldots, S$ as well as in the fusion node F in order to recover a fused estimate based on a BLUE criterion.

In order to ensure the well-posedness of the state estimation problem, the following preliminary assumptions are needed.

A6. The process noise $\mathbf{w}_k$ and the measurement noises $\mathbf{v}_k^i$, $i = 1, \ldots, S$, are zero-mean white stochastic processes with $\mathbb{E}\{\mathbf{w}_k \mathbf{w}_k'\} = \mathbf{Q}$, $\mathbb{E}\{\mathbf{v}_k^i \mathbf{w}_k'\} = \mathbf{0}$, and $\mathbb{E}\{\mathbf{v}_k^i (\mathbf{v}_k^j)'\} = \mathbf{R}^{i,j}$.

A7. $(\mathbf{A}, \mathbf{Q}^{1/2})$ is stabilizable, $(\mathbf{A}, \mathbf{C}^i)$ is detectable for $i = 1, \ldots, S$ and $\mathbf{R}^{i,i} > 0$ for $i = 1, \ldots, S$.

A. Sensor node operation

To enable the estimate transmission option, it is assumed that each sensor i has enough processing capabilities to compute some state estimate $\hat{\mathbf{x}}_{k|k}^i$ on the basis of the measurements $\mathbf{y}_{0:k}^i$ gathered at time k. More specifically, it is supposed that the estimates $\hat{\mathbf{x}}_{k|k}^i$ are recursively computed as

$$\hat{\mathbf{x}}_{k|k}^i = \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1}^i + \mathbf{K}^i \left(\mathbf{y}_k^i - \mathbf{C}^i \mathbf{A} \hat{\mathbf{x}}_{k-1|k-1}^i \right)$$

where $\mathbf{K}^i$ is the steady-state Kalman gain. As well known, under assumption A7, this is the best choice in the mean square sense among all linear time-invariant filters. Recall that the gain $\mathbf{K}^i$ can be obtained as

$$\mathbf{K}^i = \mathbf{P}^i (\mathbf{C}^i)' \left(\mathbf{R}^{i,i} + \mathbf{C}^i \mathbf{P}^i (\mathbf{C}^i)' \right)^{-1}$$

where $\mathbf{P}^i$ is the unique positive definite solution of the algebraic Riccati equation

$$\mathbf{P}^i = \mathbf{A} \left[\mathbf{P}^i - \mathbf{P}^i (\mathbf{C}^i)' \left(\mathbf{R}^{i,i} + \mathbf{C}^i \mathbf{P}^i (\mathbf{C}^i)' \right)^{-1} \mathbf{C}^i \mathbf{P}^i \right] \mathbf{A}' + \mathbf{Q} \tag{V.1}$$

B. Fusion node operation

For the fusion node F, the available information at time k is given by the vector of elapsed times $\mathbf{n}_k$ and the estimates

$$\hat{\mathbf{x}}_{k|k-n_k^i}^i = \mathbf{A}^{n_k^i} \hat{\mathbf{x}}_{k-n_k^i|k-n_k^i}^i$$

for $i = 1, \ldots, S$. Following [11], one can interpret each estimate $\hat{\mathbf{x}}_{k|k-n_k^i}^i$ as a measurement of the true state $\mathbf{x}_k$. In this context the estimation error $\hat{\mathbf{x}}_{k|k-n_k^i}^i - \mathbf{x}_k$ plays the role of a (virtual) measurement noise. With this respect, by defining

$$\mathbf{z}_k \triangleq \begin{bmatrix} \hat{\mathbf{x}}_{k|k-n_k^1}^1 \\ \vdots \\ \hat{\mathbf{x}}_{k|k-n_k^S}^S \end{bmatrix}, \ \boldsymbol{\eta}_k \triangleq \begin{bmatrix} \hat{\mathbf{x}}_{k|k-n_k^1}^1 - \mathbf{x}_k \\ \vdots \\ \hat{\mathbf{x}}_{k|k-n_k^S}^S - \mathbf{x}_k \end{bmatrix}, \ \mathbf{H} \triangleq \begin{bmatrix} \mathbf{I} \\ \vdots \\ \mathbf{I} \end{bmatrix}$$

the information available at node F can be treated as originating from the measurement channel

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \boldsymbol{\eta}_k \, .$$

In the following section it will be shown that the covariance of the virtual measurement noise $\boldsymbol{\eta}_k$ can be recursively computed as a function of $\mathbf{n}_k$, i.e.,

$$\mathbb{E}\left\{\boldsymbol{\eta}_k \boldsymbol{\eta}_k'\right\} = \boldsymbol{\Sigma}_k(\mathbf{n}_k).$$

Then the fused estimate $\hat{\mathbf{x}}_{k|k}$ can be obtained from $\mathbf{z}_k$ according to a BLUE criterion as

$$\hat{\mathbf{x}}_{k|k} = \left(\mathbf{H}' \boldsymbol{\Sigma}_k(\mathbf{n}_k)^{-1} \mathbf{H} \right)^{-1} \mathbf{H}' \boldsymbol{\Sigma}_k(\mathbf{n}_k)^{-1} \mathbf{z}_k$$

and the covariance $\mathbf{P}_{k|k}(\mathbf{n}_k)$ of the estimation error turns out to be

$$\mathbf{P}_{k|k}(\mathbf{n}_k) = \left(\mathbf{H}' \boldsymbol{\Sigma}_k(\mathbf{n}_k)^{-1} \mathbf{H} \right)^{-1} . \tag{V.2}$$

Notice that the proposed fusion rule satisfies assumptions A2 and A3. In section V-D, it will be shown that also A4 holds.

C. Computation of the virtual noise covariance

In order to derive a recursive formula for the computation of the covariance $\mathbf{\Sigma}_k(\mathbf{n}_k)$, it is convenient to decompose it as

$$\mathbf{\Sigma}_k(\mathbf{n}_k) = \begin{bmatrix} \mathbf{P}_k^{1,1}(n_k^1, n_k^1) & \cdots & \mathbf{P}_k^{1,S}(n_k^1, n_k^S) \\ \vdots & \ddots & \vdots \\ \mathbf{P}_k^{S,1}(n_k^S, n_k^1) & \cdots & \mathbf{P}_k^{S,S}(n_k^S, n_k^S) \end{bmatrix}$$

where

$$\mathbf{P}_k^{i,j}(n_k^i, n_k^j) \triangleq \mathbb{E}\left\{(\hat{\mathbf{x}}^i_{k|k-n_k^i} - \mathbf{x}_k)(\hat{\mathbf{x}}^j_{k|k-n_k^j} - \mathbf{x}_k)'\right\} .$$

Let us first focus on the case $n^i_{k+1} = n^j_{k+1} = 0$. To this end, notice that the estimation error $\hat{\mathbf{x}}^i_{k+1|k+1} - \mathbf{x}_{k+1}$ can be expressed in terms of $\hat{\mathbf{x}}^i_{k|k} - \mathbf{x}_k$ as

$$\hat{\mathbf{x}}^i_{k+1|k+1} - \mathbf{x}_{k+1} = (\mathbf{I} - \mathbf{K}^i\mathbf{C}^i)[\mathbf{A}(\hat{\mathbf{x}}^i_{k|k} - \mathbf{x}_k) - \mathbf{w}_k] + \mathbf{K}^i\mathbf{v}^i_{k+1}.$$

Then, the covariance $\mathbf{P}^{i,j}_{k+1}(0,0)$ can be obtained from $\mathbf{P}^{i,j}_k(0,0)$ by means of the recursion

$$\begin{aligned}\mathbf{P}^{i,j}_{k+1}(0,0) = & (\mathbf{I} - \mathbf{K}^i\mathbf{C}^i)(\mathbf{A}\mathbf{P}^{i,j}_k(0,0)\mathbf{A}' + \mathbf{Q})(\mathbf{I} - \mathbf{K}^j\mathbf{C}^j)' \\ & + \mathbf{K}^i\mathbf{R}^{i,j}(\mathbf{K}^j)' .\end{aligned} \tag{V.3}$$

Let us now consider generic elapsed times n^i and n^j where, without loss of generality, it is assumed that $n^i \geq n^j$. It is immediate to see that $\mathbf{P}^{i,j}_k(n^i, n^j)$ can be readily obtained from $\mathbf{P}^{i,j}_{k-n^i}(0,0)$ by propagating

$$\mathbf{P}^{i,j}_{k-n^i+l}(l,0) = (\mathbf{A}\mathbf{P}^{i,j}_{k-n^i+l-1}(l-1,0)\mathbf{A}' + \mathbf{Q})(\mathbf{I} - \mathbf{K}^j\mathbf{C}^j)' \tag{V.4}$$

for $l = 1, \ldots, n^i - n^j$ and then propagating

$$\begin{aligned}&\mathbf{P}^{i,j}_{k-n^j+l}(n^i - n^j + l, l) = \\ &\quad \mathbf{A}\mathbf{P}^{i,j}_{k-n^i+l-1}(n^i - n^j + l - 1, l)\mathbf{A}' + \mathbf{Q}\end{aligned} \tag{V.5}$$

for $l = 1, \ldots, n^j$. Hereafter, for the sake of compactness, the above described propagation will be written as

$$\mathbf{P}^{i,j}_k(n^i, n^j) = \mathbf{f}^{i,j}_{n^i,n^j}\left[\mathbf{P}^{i,j}_{k-n^i}(0,0)\right] .$$

D. Asymptotic behavior of the virtual noise covariance

Once again, let us first consider the case $\mathbf{n} = \mathbf{0}$. The asymptotic behavior of the covariance $\mathbf{\Sigma}_k(\mathbf{0})$ can be analyzed by noting that the recursions (V.3) are equivalent to

$$\mathbf{\Sigma}_{k+1}(\mathbf{0}) = \mathbf{\Psi}\mathbf{\Sigma}_k(\mathbf{0})\mathbf{\Psi}' + \mathbf{\Omega}$$

where

$$\mathbf{\Psi} \triangleq \mathrm{diag}\left(\mathbf{A} - \mathbf{K}^1\mathbf{C}^1\mathbf{A}, \ldots, \mathbf{A} - \mathbf{K}^S\mathbf{C}^S\mathbf{A}\right)$$

and $\mathbf{\Omega}$ is a suitable matrix that accounts for the terms dependent on $\mathbf{Q}$ and $\mathbf{R}^{i,j}$. Under assumption A7, the matrix $\mathbf{\Psi}$ is Schur stable and the matrix $\mathbf{\Omega}$ is positive definite, then as k tends to infinity the covariance $\mathbf{\Sigma}_k(\mathbf{0})$ exponentially converges to the unique positive definite solution $\mathbf{\Sigma}(\mathbf{0})$ of the Lyapunov equation

$$\mathbf{\Sigma}(\mathbf{0}) = \mathbf{\Psi}\mathbf{\Sigma}(\mathbf{0})\mathbf{\Psi}' + \mathbf{\Omega} .$$

Let $\mathbf{P}^{i,j}(0,0)$ denote the steady-state value of $\mathbf{P}^{i,j}_k(0,0)$ obtained from the corresponding block of $\mathbf{\Sigma}(\mathbf{0})$. Further, let us consider generic elapsed times n^i and n^j with $n^i > n^j$. Since at each time instant $\mathbf{P}^{i,j}_k(n^i, n^j)$ can be obtained from $\mathbf{P}^{i,j}_{k-n^i}(0,0)$ through the continuous function $\mathbf{f}^{i,j}_{n^i,n^j}(\cdot)$, then it follows that $\mathbf{P}^{i,j}_k(n^i, n^j)$ converges exponentially to the steady-state value

$$\mathbf{P}^{i,j}(n^i, n^j) \triangleq \mathbf{f}^{i,j}_{n^i,n^j}\left(\mathbf{P}^{i,j}(0,0)\right) .$$

Finally, it is important to notice that the foregoing discussion implies that also assumption A4 holds for the considered fusion rule since the covariance $\mathbf{P}_{k|k}(\mathbf{n})$ is a continuous function of $\mathbf{\Sigma}_{k|k}(\mathbf{n})$ (see (V.2)).

VI. Conclusions

The paper has addressed the state estimation problem in a centralized sensor network assuming that: (1) estimates are required at a distant location from the sensors connected via a communication link; (2) a limitation on the communication rate of each sensor is imposed; (3) each sensor node has enough processing capability to compute local state estimates. Measurement-independent strategies for deciding which data transmit have been investigated. It has been shown that: (1) boundedness of the mean square error is guaranteed provided that an upper bound on the inter-transmission time is enforced; (2) communication strategies with optimal performance in terms of minimum mean square error can be devised. Future work will concern the extension to measurement-dependent transmission strategies and distributed sensor networks.

References

[1] X. Li and W.S. Wong: "State estimation with communication constraints", *Systems & Control Letters*, vol. 28, pp. 49-54, 1996.

[2] W.S. Wong and R.W. Brockett: "Systems with finite communication bandwidth constraints - Part I: state estimation problem", *IEEE Trans. on Automatic Control*, vol. 42, pp. 1294-1299, 1997.

[3] X. Li and W.S. Wong: "Constrained state estimation for systems with finite communcation bandwidth", *Proc. 37th IEEE Conf. on Decision and Control*, pp. 257-262, Tampa, USA, 1998.

[4] G.N. Nair and R.J. Evans: "State estimation under bit-rate constraints", *Proc. 37th IEEE Conf. on Decision and Control*, pp. 251-256, Tampa, USA, 1998.

[5] G.N. Nair and R.J. Evans: "A finite-dimensional coder-estimator for rate-constrained state estimation", *Proc. of the 14th IFAC World Congress*, pp. 19-24, Beijing, China, 1999.

[6] G.N. Nair and R.J. Evans: "Structural results for finite bit-rate state estimation", *Proc. 38th IEEE Conf. on Decision and Control*, pp. 47-52, Adelaide, Australia, 1999.

[7] V. Gupta, B. Hassibi, and M. Murray, "Optimal LQG control across packet-dropping links", *System & Control Letters*, vol. 56, pp. 439-446, 2007.

[8] G. Battistelli, A. Benavoli, and L. Chisci, "State estimation with a remote sensor under limited communication rate", *Proc. 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP 2008)*, St. Julians, Malta, 2008.

[9] N. K. Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica*, vol. 26, pp. 373395, 1984.

[10] I. Adler and N. Megiddo, "A simplex algorithm whose average number of steps is bounded between two quadratic functions of the smaller dimension," *Journal of the Association for Computing Machinery*, vol. 32, pp. 871895, 1985.

[11] X. R. Li, Y.-M. Zhu, J. Wang, and C.-Z. Han, "Optimal linear estimation fusion-part I: unified fusion rules *IEEE Transactions on Information Theory*, vol. 49, pp. 2192-2208, 2003.